

UNIVERSIDAD AUTONOMA DE MADRID

ESCUELA POLITECNICA SUPERIOR



Grado en Ingeniería Informática

TRABAJO FIN DE GRADO

**SISTEMA INFORMÁTICO DE APOYO A LAS ANALÍTICAS
PARA EL APRENDIZAJE (LEARNING ANALYTICS) PARA
ENTORNOS EDUCATIVOS ON-LINE**

Juan Soberón Higuera
Tutor: Ruth Cobos Pérez

JULIO 2020

SISTEMA INFORMÁTICO DE APOYO A LAS ANALÍTICAS PARA EL APRENDIZAJE (LEARNING ANALYTICS) PARA ENTORNOS EDUCATIVOS ON-LINE

AUTOR: Juan Soberón Higuera
TUTOR: Ruth Cobos Pérez

Dpto. Ingeniería Informática
Escuela Politécnica Superior
Universidad Autónoma de Madrid
julio de 2020

Resumen

Los cursos MOOC (*Massive Open Online Courses*) son una herramienta de aprendizaje en constante crecimiento en oferta (número de cursos ofertados), demanda (número de estudiantes matriculados) y relevancia en la disciplina del aprendizaje en línea. En este contexto, se han detectado potenciales factores relacionados con la insuficiente interacción estudiante-profesor y el aislamiento que sienten los estudiantes que pueden afectarles negativamente y que precisan de un amplio estudio para ser analizados, comprendidos y, si se requiere, mitigados o solucionados. Si este objetivo se cumple, se habrá conseguido dar un gran paso para hacer del aprendizaje con este tipo de recursos un proceso eficiente y útil para cualquier estudiante que desee utilizarlos. Además, los MOOC recogen gran cantidad de información en relación con la interactividad del estudiante con sus recursos, con lo que son una gran fuente de datos en este campo. La disciplina que facilita poder afrontar el problema planteado es la Analítica del Aprendizaje o *Learning Analytics*.

El presente Trabajo de Fin de Grado busca avanzar dentro de la comprensión de este tipo de sensaciones. Para ello, se propone el desarrollo, implementación y explotación de una plataforma escalable (*edX-LIMS: Learning Intervention Monitoring Service for edX MOOCs*) que permita realizar un proceso de acompañamiento de los estudiantes de un MOOC. Dicha plataforma proporciona periódicamente a los estudiantes de un MOOC información visual en un *Dashboard* o Panel de aprendizaje en la Web, mostrándoles su progreso y participación en el MOOC. Esta información proporcionada es parte de una estrategia de intervención sobre el aprendizaje de estos estudiantes. El sistema ofrece también a los instructores de MOOC acceso a un *Dashboard* o Panel de Instructores en la Web que muestra el interés en este servicio por parte de los estudiantes y, por lo tanto, facilita la evaluación del éxito o el fracaso de la estrategia de intervención.

Abstract

Massive Open Online Courses, also known as MOOCs, are a learning tool which are continuously growing in terms of supply (number of courses offered), demand (number of learners enrolled) and relevance in e-Learning discipline. Regarding this, some factors related to the insufficient student-teacher interaction and the isolation felt by the students have been detected during the course development. These factors could negatively affect them and require a comprehensive study to be analyzed, understood, and, if required, mitigated or solved. If this objective is achieved, it will be one step closer to make learning using this type of resources an efficient and useful process for any learner who wishes to do it. In addition, MOOCs collect large amounts of information regarding learner interactivity with the course resources, which makes MOOCs a great source of data. The discipline that makes it easier to face the problem posed is Learning Analytics.

This End-of-Grade Study seeks to advance within the understanding of this type of feeling. In order to do it, it proposes the development, implementation and exploitation of a scalable platform (*edX-LIMS: Learning Intervention Monitoring Service for edX MOOCs*) which will be able to give feedback to MOOC learners about its course performance by periodically providing them with visual information on a Dashboard or Learner Web Panel showing them their progress and participation in the MOOC. This information provided is part of an intervention strategy on the learning of these learners. The system also offers

MOOC instructors access to a Dashboard or Instructor Web Panel that shows learners' interest in this service and, therefore, helps to evaluate the success or failure of the intervention procedure.

Palabras clave

Cursos Abiertos Masivos Online, MOOC, Analítica del Aprendizaje, Intervención, Estrategia de Aprendizaje, Retroalimentación, Implicación, Tecnología Web

Keywords

Massive Open Online Courses, MOOC, Learning Analytics, Intervention, Learning Strategy, Feedback, Engagement, Web-based Technology.

Agradecimientos

A Ruth Cobos, por la oportunidad de realizar este proyecto y por el infinito apoyo y ayuda a todos los niveles que me ha brindado durante todo este tiempo.

A Idoia Alarcón y a la Administración de la EPS, por poner todo de su parte para que este Trabajo de Fin de Grado llegara a buen puerto lo antes posible. A Alberto Blázquez, Juan Carlos Ruíz y Sorin Peia, por su ayuda y colaboración durante las distintas fases de este Trabajo.

A toda mi familia, por todo el apoyo incondicional recibido durante demasiados años y estar siempre ahí pasara lo que pasara. No tengo palabras suficientes de agradecimiento.

A Bea, por el apoyo, la paciencia y el cariño a cambio de nada que no has dejado de darme durante todos estos meses de trabajo. A Alberto y Pablo, por darme un techo y no retractarse (de momento). A Chusco, Jose, Manu, Nico y el resto de los morenos por ser los mejores acompañantes con los que hacer este camino. A Héctor por los cafés y las cervezas de desahogo en la cafetería de la EPS o donde fuera. A Bego, Cholfo, Cuesta, Tomy y asociados por ser mi refugio en el norte. A Huesca y Sara, porque esto de alguna manera también es suyo.

INDICE DE CONTENIDOS

1	Introducción.....	1
1.1	Motivación.....	1
1.2	Objetivos.....	1
1.3	Estructura del documento	2
2	Estado del arte	5
2.1	Análisis del aprendizaje online.....	5
2.2	Trabajo previo: herramientas de análisis de MOOCs.....	6
2.2.1	edX-MAS y edX-MAS+.....	6
2.2.2	edX-WS	6
2.3	Análisis del contexto: edX-LIS	7
2.3.1	Arquitectura lógica y módulos	7
2.3.2	Lenguajes de desarrollo	8
2.4	Estudio de tecnologías	9
2.4.1	Modelos de bases de datos.....	9
2.4.2	Lenguajes de desarrollo y librerías	10
2.4.1	Impacto y satisfacción	11
3	Diseño.....	12
3.1	Descripción del sistema.....	12
3.1.1	Servicio de procesamiento de datos del curso	13
3.1.2	Servicio de intervención en el aprendizaje	15
3.1.3	Servicio de visualización de intervención	16
3.1.4	Servicio de monitorización de datos del curso	16
3.1.5	Servicio de monitorización de intervención	17
3.2	Diseño del sistema.....	17
3.2.1	Servicio de procesamiento de datos del curso	17
3.2.2	Servicio de intervención en el aprendizaje	23
3.2.3	Servicio de visualización de intervención	26
3.2.4	Servicio de monitorización de datos del curso	27
3.2.5	Servicio de visualización de intervención	28
4	Desarrollo, implementación y pruebas	29
4.1	Estructura de ficheros	29
4.2	Implementación y escalabilidad	30
4.3	Pruebas	30
4.3.1	Pruebas unitarias.....	31
4.3.2	Pruebas de integración y funcionamiento.....	31
4.3.3	Pruebas de sistema.....	32
5	Validación y resultados	33
5.1	Valoración de la estrategia de intervención en el aprendizaje.....	33
5.2	Valoración de la monitorización de la estrategia de intervención.....	36
6	Conclusiones y trabajo futuro.....	37
6.1	Conclusiones.....	37
6.2	Trabajo futuro	38
	Referencias	39
	Anexos.....	I
A	Encuesta de satisfacción con edX-LIS	I
B	Desglose de eventos para la generación de indicadores	IV
C	Vistas de edX-LIMS	VII

INDICE DE FIGURAS

FIGURA 3-1: SERVICIOS DE LA PLATAFORMA EDX-LIMS.....	12
FIGURA 3-2: DEFINICIÓN DEL SERVICIO DE PROCESADO DE DATOS DEL CURSO	15
FIGURA 3-3: DEFINICIÓN DE LA CLASE LOGPROCESSOR	18
FIGURA 3-4: MODELO DE DATOS DE LA PLATAFORMA EDX-LIMS.....	22
FIGURA 3-5: DEFINICIÓN DE LA CLASE EDXMONGOSTORE.....	23
FIGURA 3-6: ESQUEMA DEL MODELO VISTA CONTROLADOR	24
FIGURA 3-7: ESTRUCTURA DE NAVEGACIÓN DE EDX-LIMS.....	24
FIGURA 3-8: DEFINICIÓN DE LOS MÉTODOS DE <i>MAILING.PY</i>	25
FIGURA 3-9: DEFINICIÓN DE LOS MÉTODOS DE <i>USER_DASHBOARD.PY</i>	27
FIGURA 3-10: VISTA DE <i>DASHBOARD</i> DE PROFESOR.....	27
FIGURA 3-11: VISTA DEL SERVICIO DE VISUALIZACIÓN DE INTERVENCIÓN	28
FIGURA 5-1: RESPUESTA A PREGUNTAS SOBRE LA PERCEPCIÓN DE LOS ESTUDIANTES SOBRE SU ACOMPañAMIENTO.....	34
FIGURA 5-2: RESPUESTA A PREGUNTAS SOBRE LA OPINIÓN DE LOS ESTUDIANTES ACERCA DE SU <i>DASHBOARD</i>	34
FIGURA 5-3: RESPUESTA A PREGUNTAS SOBRE LA PERCEPCIÓN DE LOS ESTUDIANTES ACERCA DE LOS BENEFICIOS DE UN <i>DASHBOARD</i> DE ACOMPañAMIENTO.....	35
FIGURA 5-4: RESPUESTA A PREGUNTAS SOBRE LA COMPRENSIÓN DE LOS ESTUDIANTES DE LA INFORMACIÓN DE SU <i>DASHBOARD</i> DE ACOMPañAMIENTO	35
FIGURA A-1: IMAGEN DE <i>DASHBOARD</i> DE EDX-LIS UTILIZADA EN LA ENCUESTA	II
FIGURA C-1: GRÁFICO INDICADOR DE NOTA ACTUAL DE USUARIO	VII
FIGURA C-2: GRÁFICOS DE LÍNEA CON LOS VALORES DE LOS INDICADORES DEL USUARIO	VII
FIGURA C-3: GRÁFICOS DE RADAR Y DE BARRA CON EL PROGRESO POR UNIDAD DEL ESTUDIANTE	VIII

FIGURA C-4: VISTA DE <i>DASHBOARD</i> DE ESTUDIANTE	VIII
FIGURA C-5: VISTA DE <i>/MAILING</i>	IX
FIGURA C-6: DETALLE DE SELECCIÓN DE <i>TEMPLATE</i> DE EMAIL	X
FIGURA C-7: DETALLE DE SELECCIÓN DE DESTINATARIOS DE EMAIL	X
FIGURA C-8: DETALLE DE CONFIRMACIÓN DE CONTENIDO DE EMAIL	X
FIGURA C-9: VISTA DE HISTORIAL DE ENVÍOS PREVIOS	X
FIGURA C-10: VISTA DE <i>DASHBOARD</i> DE PROFESOR	XI
FIGURA D-1: PROGRAMA DE LASI 2020	XII

INDICE DE TABLAS

TABLA 2-1: ARCHIVOS DE DATOS E INDICADORES GENERADOS POR EDX-LIS	8
TABLA 3-1: TABLA DE INDICADORES DE EDX-LIMS	14
TABLA 3-2: TABLA DE COLECCIONES DE EDX-LIMS	19
TABLA 4-1: ESTRUCTURA DE FICHEROS DE /	29
TABLA 4-2: ESTRUCTURA DE FICHEROS DE /ASSETS	29
TABLA 4-3: ESTRUCTURA DE FICHEROS DE /STORAGE	29
TABLA B-1: EVENTOS DE VÍDEO DE EDX TOMADOS EN CUENTA EN EDX-LIMS	IV
TABLA B-2: EVENTOS DE PROBLEMA DE EDX TOMADOS EN CUENTA EN EDX-LIMS	V
TABLA B-3: EVENTOS DE NAVEGACIÓN DE EDX TOMADOS EN CUENTA EN EDX-LIMS	V
TABLA B-4: EVENTOS DE FORO DE EDX TOMADOS EN CUENTA EN EDX-LIMS	VI

1 Introducción

El siguiente capítulo define cuales son la motivación, los objetivos y la estructura del presente Trabajo de Fin de Grado (TFG).

1.1 Motivación

En los últimos años, los modelos de aprendizaje autorregulado han proliferado en el marco del aprendizaje en línea. De entre ellos destacan especialmente los MOOC (acrónimo de *Massive Online Open Courses*, Cursos Online Masivos y Abiertos), cursos masivos online en los que el estudiante tiene todos los recursos disponibles para su realización, pero debe planificar su propio proceso de aprendizaje y todas las acciones asociadas a él. La naturaleza del modo de impartición del curso, en el que figuras tradicionales como la presencia de un tutor que ejerza como elemento de control y validación del trabajo del estudiante no existe, hace que en muchas ocasiones la relación de un alumno con su propio proceso de aprendizaje tenga particularidades especiales. Concretamente, más allá de las destrezas necesarias para comprender y superar los objetivos de un curso, en el estudiante influyen un nuevo tipo de sensaciones asociadas a la falta de validación o reconocimiento de su trabajo que pueden afectar negativamente a su progreso en el curso, o llevarle incluso a su abandono.

Dentro del ecosistema de MOOCs, la Universidad Autónoma de Madrid lleva más de 5 años ofreciendo cursos de este tipo encuadrados en la plataforma de educación en línea del consorcio edX [37]. La participación de los estudiantes en los numerosos cursos genera una gran cantidad de información acerca de su actividad durante el tiempo que dedican a su realización, que resultan cruciales para el desarrollo de estudios y herramientas destinadas a su análisis [19][24][33]. La última de estas herramientas es edX-LIS, cuyo objetivo es la creación de una estrategia de acompañamiento al estudiante para intentar mantener su compromiso con el curso.

De la experiencia adquirida durante el desarrollo de edX-LIS surge la necesidad de crear una plataforma que procure un marco para este tipo de herramientas y habilite un servicio de uso ambivalente; una aplicación web que permita por un lado el uso autónomo del profesor para ayudarle con la estrategia de acompañamiento tanto en la recogida y representación de datos como en crear la generación de información adicional que le sea de utilidad, y que al mismo tiempo proporcione al estudiante los recursos necesarios para sentirse acompañado durante la realización del curso.

1.2 Objetivos

El objetivo del presente Trabajo de Fin de Grado es el desarrollo de una plataforma llamada edX-LIMS (*Learning Intervention Monitoring Service for edX MOOCs*) que permita tanto ejercer como extender la funcionalidad de herramientas previamente existentes creadas para analizar y realizar el proceso de acompañamiento de los usuarios de un curso. Para cumplir este propósito, se plantea la consecución de sucesivos hitos que lleven a su creación y validación final:

- Analizar el estado actual de los MOOCs.

- Evaluar y analizar cuáles son los principales motivos que afectan al aprendizaje y éxito del usuario del curso, y las ventajas y beneficios del uso de sistemas de acompañamiento y *feedback* a usuarios de cursos online.
- Evaluar el trabajo previo realizado en el ámbito de la analítica de aprendizaje y las soluciones concretas que se han utilizado o se están utilizando actualmente en relación con ello.
- Evaluar la satisfacción actual, la percepción y las necesidades que los estudiantes de un MOOC puedan expresar acerca de mecanismos de acompañamiento actualmente utilizados.
- Estudiar las tecnologías adecuadas para la realización del proyecto.
- Diseñar la plataforma. Este proceso incluye diversos sub-hitos:
 - Plantear un modelo de base de datos adecuado.
 - Diseñar las clases que conformen la plataforma.
 - Diseñar las vistas que compondrán la plataforma y definir los datos a mostrar en ellas.
- Desarrollar la plataforma. Esto incluye también varios sub-hitos
 - Importar datos de un MOOC y generar indicadores en relación con los eventos de actividad registrados de los estudiantes del curso.
 - Almacenar la información obtenida de acuerdo con el modelo de base de datos definido.
 - Implementar la visualización de datos (*Dashboard*) de un usuario que presente datos que él considere relevantes.
 - Implementar la visualización de datos (*Dashboard*) de un curso que cumpla con lo necesario para que un profesor pueda seguir la evolución de sus estudiantes.
 - Crear un mecanismo que permita conocer en más detalle cuál es la relación entre el estudiante y el mecanismo de acompañamiento que se le proporciona.
 - Crear un servicio que gestione el envío de credenciales de acceso a la plataforma de los usuarios.
 - Realizar las pruebas necesarias sobre todos los elementos que conforman la plataforma.
- Evaluar si la implementación de la nueva plataforma supone una mejora en comparación con las herramientas pasadas tanto para el estudiante como para el profesor del curso.
- Plantear posibles mejoras o evoluciones a implementar en un futuro.

1.3 Estructura del documento

La presente memoria se compone de los siguientes capítulos:

- El *primer capítulo* trata la motivación, la propuesta y los objetivos del Trabajo.
- El *segundo capítulo* presenta el Estado del Arte, en donde se realiza un análisis de las distintas herramientas que se han ido utilizando en el marco de la investigación acerca de la analítica de aprendizaje, y se lleva a cabo un análisis más en detalle de una de ellas en concreto, edX-LIS. También incluye un análisis sobre las tecnologías a utilizar.
- El *tercer capítulo* explica la funcionalidad y el diseño de la plataforma mediante su división en los distintos servicios que la componen.

- El *cuarto capítulo* expone cual ha sido su estructura de ficheros para su despliegue, el análisis de hipotéticas implementaciones y las pruebas llevadas a cabo sobre el sistema.
- El *quinto capítulo* explica las valoraciones y resultados de la experiencia con los servicios de la plataforma por parte de estudiantes y profesores del curso.
- El *sexto capítulo* presenta finalmente las conclusiones del Trabajo y el análisis del trabajo futuro a desarrollar.

De manera adicional, se incluyen los siguientes anexos:

- El *Anexo A* contiene información acerca de la encuesta de satisfacción llevada a cabo sobre estudiantes de un curso MOOC con relación al sistema de acompañamiento de uso previo a edX-LIMS.
- El *Anexo B* contiene información acerca de la estructura de eventos de logs de un curso sobre los que se han construido los indicadores de interacción con un MOOC tomados en cuenta por edX-LIMS.
- El *Anexo C* contiene las distintas vistas que componen la plataforma en varios de sus servicios.
- El *Anexo D* contiene el artículo presentado en el LASI 2020 (*Learning Analytics Summer Institute Spain 2020*) [18] en junio de ese año, evento de referencia en el área de las analíticas para el aprendizaje, organizado por SNOLA (*Spanish Network Of Learning Analytics*), en colaboración con SOLAR (*Society for Learning Analytics Research*) [36].

2 Estado del arte

Este capítulo analizará el estado del arte del aprendizaje online en el contexto de los cursos masivos online o MOOCs, los trabajos previos realizados en este campo que han sido la base para el presente TFG y las tecnologías utilizadas y que van a utilizarse en él.

2.1 Análisis del aprendizaje online

Los MOOC son cursos online a los que cualquier persona puede acceder y cursar. Estos cursos son habitualmente creados e impartidos por distintas universidades, que a su vez suelen publicarlos a través de plataformas que aglutinan las múltiples fuentes en un marco único. Según el tipo de curso pueden existir varias opciones de cursado, siendo la fórmula más habitual la opción de acceso gratuito y limitado en el tiempo a los contenidos del curso, y una opción de pago (que se denomina itinerario verificado en muchos entornos) para poder acceder a la evaluación del curso y la posibilidad de conseguir un certificado que acredite haberlo completado emitido por la entidad creadora del curso.

Desde la aparición del primer MOOC en 2011, la evolución en su uso ha sido constante. En su octavo año de vida como tal, los MOOCs siguen creciendo en oferta y demanda: se calcula que en 2019 más de 110 millones de estudiantes han cursado 13500 cursos de más de 900 universidades con programa de MOOC [3]. Hoy en día, los MOOCs constituyen un pilar en la formación libre y el desarrollo de competencias en numerosos campos, y todo apunta a que su rol e importancia sólo irán en aumento con el paso del tiempo.

Para el análisis de la relación entre un MOOC y un usuario/estudiante de este, es necesario un enfoque que permita dar respuesta al origen de los datos obtenidos a lo largo del proceso de realización del curso (resultados, abandono del curso, etcétera). La comprensión del origen y el contexto de estos datos permitirá la optimización de los contenidos, recursos y herramientas destinadas a facilitar que el seguimiento de un curso por parte de un usuario llegue a ser un éxito (esto es, que se cumpla tanto que no lo abandone a la mitad como que supere la nota necesaria para obtener su acreditación). El área de conocimiento que lleva a cabo estas tareas es la analítica del aprendizaje o Learning Analytics [34].

Las analíticas del aprendizaje pueden ser de varios tipos:

- **Analítica descriptiva:** consiste en los procesos de recopilación y agregación de datos (*logs*) de un curso generados por la actividad de los estudiantes del curso tanto con los contenidos de este como entre ellos, y la generación de información directamente relacionada con esas interacciones.
- **Analítica predictiva:** se trata de la analítica dirigida al desarrollo de modelos estadísticos y aprendizaje automático para generar una predicción de posibles situaciones futuras con los indicadores actuales. Con ello, se busca poder adelantarse a una potencial situación negativa (por ejemplo, el abandono de un curso o una mal resultado al final de este) con el fin de poder tomar medidas antes de que esto ocurra.
- **Analítica prescriptiva:** con la información tomada en las dos analíticas anteriores, la analítica prescriptiva busca el establecimiento de procesos y funcionalidades que permitan interactuar con el estudiante para ofrecerle un servicio de apoyo que evite que potenciales eventos negativos lleguen a ocurrir.

En la relación entre este tipo de analíticas y los MOOCs, se ha detectado que diversos factores afectan especialmente a los niveles de retención de estudiantes en el curso (esto es, estudiantes que no abandonan el curso) y de actividad y participación con sus contenidos y recursos. Algunos de estos factores son los siguientes:

- **Retroalimentación insuficiente:** aunque el estudiante realice las actividades y problemas propuestos en el curso, en muchas ocasiones recibe un *feedback* escaso o hasta inexistente. Esto le provoca una sensación de desánimo al no percibir que su esfuerzo sea reconocido y valorado.
- **Falta de interacción entre profesores y estudiantes:** si el profesor del curso no participa dinámicamente con el usuario o se muestra ausente a la hora de interactuar con los estudiantes, estos se sienten en muchas ocasiones abandonados, lo cual repercute negativamente en el interés/motivación que pasan a dedicar al curso.

Sensación de aislamiento: directamente derivado de los factores previos, el estudiante se ve sólo a la hora de afrontar la totalidad del curso y sin un acompañamiento en su proceso de aprendizaje como potencial apoyo, que bien podría ser de tipo tanto personal (por parte de los profesores) como visual (*feedback* de su progreso durante el curso). Ante esta situación, son muchos los estudiantes que, desmotivados, deciden abandonar el curso.

2.2 Trabajo previo: herramientas de análisis de MOOCs

El Grupo de Herramientas Interactivas Avanzadas (GHIA, vghia.ii.uam.es/) de la UAM, al que pertenece la tutora del presente Trabajo de Fin de Grado, tiene entre sus líneas de investigación el análisis de datos en entornos educativos y cursos en línea (Learning Analytics). La UAM pertenece al consorcio edX desde 2014, y muchos de los trabajos realizados tienen como materia de análisis los datos pertenecientes a esos cursos propios ofertados en edX. Con ese fin se han desarrollado varias herramientas dirigidas a múltiples campos de estudio que se presentan en los próximos subapartados [1]. Además, es de destacar que la comunidad española que trabaja en Learning Analytics ofrece también gran variedad de herramientas en esta línea [20].

2.2.1 edX-MAS y edX-MAS+

El sistema edX-MAS [19][5] (acrónimo de *Model Analyzer System for edX MOOCs*) es una herramienta desarrollada con el objetivo de crear modelos predictivos a partir de la importación de datos de uso de un curso MOOC de UAMx [38]. Con la generación de estos modelos de frecuencia diaria, se puede predecir el discurrir de un estudiante, las posibilidades que tiene de lograr la obtención del certificado de curso (aprobado). También incluye la recomendación del mejor algoritmo para generar el modelo, visualización de gráficas y exportación de resultados. Esta herramienta fue evolucionada para también predecir las posibilidades del abandono de los estudiantes del curso, así como la inclusión de nuevos indicadores y múltiples mejoras funcionales en un nuevo sistema de similar arquitectura y planteamiento llamado edX-MAS+ [24][6].

2.2.2 edX-WS

El sistema edX-WS (acrónimo de *Warning System for edX MOOCs*) es una herramienta de análisis prescriptivo que busca detectar cuáles son los estudiantes con mayor riesgo de abandono del curso en base a su actividad. Con su evaluación, el sistema genera recomendaciones a seguir para cada estudiante de cara a la mejora de sus prestaciones al

realizar el curso y crea una visualización gráfica para el seguimiento de los usuarios afectados [33].

2.3 Análisis del contexto: edX-LIS

Durante estos últimos años y tras la utilización de las herramientas antes presentadas, se detectó que en parte la gran tasa de abandono que tienen estos cursos es porque los estudiantes desarrollan una sensación de abandono mientras realizan el curso. La carencia de un *feedback* que mitigue ese aislamiento que percibe el usuario hace que este en muchas ocasiones, cansado y desmotivado, decida no continuar con el estudio del curso (véase la sección 2.1 para más información).

Para mitigar estos factores que se han identificado como principales problemas a la hora de continuar con el estudio de un MOOC, se desarrolló la herramienta edX-LIS (acrónimo de *Learning Intervention System for edX MOOCs*), la cual permite la interacción con los estudiantes del MOOC mediante el envío periódico de mensajes de *feedback* con la información de su progreso y su actividad a lo largo del curso [7].

2.3.1 Arquitectura lógica y módulos

El sistema edX-LIS se compone de tres módulos, cada uno de ellos con una funcionalidad diferenciada:

- **Importación de datos y generación de indicadores.** El sistema toma todos los ficheros de logs existentes del curso y los procesa en función del tipo de evento para generar los indicadores que miden el progreso del usuario del curso (véase Anexo B para obtener más información sobre la tipología de los eventos contemplados). Esta ingesta y procesamiento de datos genera una salida en forma de múltiples ficheros CSV, cada uno de los cuales representa un dato distinto:

Nombre fichero	Indicador
<i>aggrData</i>	Indicadores agregados por estudiante
<i>allEvents</i>	Todos los eventos registrados por estudiante
<i>allEvents_accountabilityFiltered</i>	Todos los eventos registrados por cada estudiante (excepto eventos exclusivos del itinerario verificado)
<i>allIndicators</i>	Indicadores no agregados por usuario y día
<i>distinctEvents_perUser</i>	Número de eventos de cada tipo por usuario y día
<i>Enrollment</i>	Tipo de itinerario de curso de cada estudiante
<i>events_User</i>	Total de eventos realizados por estudiante y día
<i>grades_problemInfo</i>	Nota de estudiante en cada uno de los problemas del curso
<i>sessions_User</i>	Número de sesiones realizadas por estudiante y día
<i>timeEvent_userDay</i>	Número de eventos de cada tipo realizados por estudiante y día
<i>videoInfo</i>	Identificadores de videos vistos por cada estudiante

<i>finalPData</i>	Problemas del curso y peso
<i>gradeReport</i>	Nota de cada estudiante en cada subunidad del curso
<i>gradeReport_avgs</i>	Media del estudiante en cada unidad del curso y nota final
<i>unitWeight</i>	Peso en porcentaje de la nota de cada unidad del curso
<i>passedStudents</i>	Estudiantes que han aprobado (nota por unidades y nota final)
<i>nProblems_perStudent_perUnit</i>	Para cada estudiante, problemas realizados de cada unidad del curso
<i>nProblems_perUnit</i>	Número de problemas de cada unidad del curso
<i>problemsToBeDone</i>	Problemas pendientes de realizar por cada estudiante
<i>undoneAttainableGrade</i>	Nota que aún puede obtener en cada subunidad por estudiante
<i>undoneAttainableGrade_avgs</i>	Nota que aún puede obtener en cada unidad y en total de cada estudiante
<i>successViability</i>	Posibilidad de cada estudiante de aprobar con la nota de los ejercicios que tienen pendientes de hacer
<i>cohors_file</i>	Para cada estudiante, sus indicadores de problemas, videos y nota

Tabla 2-1: Archivos de datos e indicadores generados por edX-LIS

- **Generación de información de usuario.** Con la información generada en el módulo anterior, se lleva a cabo la generación de la información visual de datos del usuario. Esto se lleva a cabo mediante la generación de gráficos como imágenes fijas.
- **Envío de emails:** Se lleva a cabo la composición de una única imagen a partir de los varios gráficos e indicadores generados en el módulo anterior. Esta imagen se inserta dentro de un email con un contenido predefinido que se envía a una selección total o parcial de estudiantes del curso. Un detalle de la imagen con los gráficos e indicadores contenidos en los emails enviados a los estudiantes se puede observar en la figura A-1.

2.3.2 Lenguajes de desarrollo

La herramienta edX-LIS está desarrollada mediante varios scripts en **Python** y **R** según el módulo de la herramienta. Ambos son lenguajes *Open Source* interpretados y multiplataforma, y ampliamente utilizados en *data science* con fines de análisis estadístico y gráfico.

Se utiliza **Python** en el **módulo de importación de datos** para la lectura de logs y su agrupación y tratamiento para la generación de varios de indicadores de actividad. También se utiliza en el **módulo de envío de emails** para la integración de imágenes y gráficos en una única imagen mediante la librería PIL [30].

Por su parte, se utiliza **R** para generar las funcionalidades del **módulo de importación de datos** relativas al cálculo de notas en unidades del curso, para todo el **módulo de generación de información de estudiante** a través del paquete *ggplot2* [40], y para el **módulo de envío de emails** mediante el paquete *mailR* [9].

El sistema no utiliza ningún tipo de base de datos, y gestiona el intercambio de información entre módulos mediante la lectura y escritura de información en los CSV descritos en la Tabla 2-1. Esto provoca que para generar la información actualizada de todos los estudiantes del curso cada vez que se ejecuta la herramienta sea necesario procesar todos los logs previos del curso.

2.4 Estudio de tecnologías

Esta sección consta del análisis de las distintas posibilidades existentes a la hora de escoger las tecnologías para llevar a cabo el presente Trabajo de Fin de Grado, y la selección final basada en el susodicho análisis. Se contempla este estudio para lenguajes de programación y bases de datos.

2.4.1 Modelos de bases de datos

Las bases de datos se clasifican en dos grandes grupos: relacionales o no relacionales.

Bases de datos relacionales

Una base de datos relacional [31] es aquella que sigue un modelo relacional, esto es, que almacena y proporciona acceso a puntos de datos relacionados entre sí. Los datos pasan a tener un identificador único y se enmarcan en agrupaciones cuya estructura de datos ha sido previamente definida llamadas tablas. Para trabajar con este tipo de bases de datos se utiliza el lenguaje SQL (*Structured Query Language*).

Los sistemas de gestión de bases de datos más extendidos en la actualidad [13] son Oracle [25], MySQL [23], Microsoft SQL Server [21] y PostgreSQL [28].

Bases de datos no relacionales

Las bases de datos no relacionales [2] son todas aquellas que no se ajustan al modelo relacional, lo que engloba en realidad a una multitud de tipos de modelo de datos. Tienen una serie de características generales comunes (por ejemplo, no existir relaciones o no tener estructuras de datos predefinidas). Su surgimiento nace de la necesidad de tratar con una cantidad cada vez mayor de datos, y de los problemas de escalabilidad de las bases de datos relacionales a la hora de tratar con ese incremento de información. Este tipo de bases de datos poseen como características generales la flexibilidad, escalabilidad y alta funcionalidad a la hora de gestionar la información.

Actualmente [13], los sistemas de gestión más utilizados en este tipo de bases de datos son MongoDB [22], Elasticsearch [15] y Cassandra [4].

Elección de modelo y sistema de gestión

Dado que no existe una tecnología previa sobre la que sustentarse, (edX-LIS no dispone de base de datos, véase apartado 2.2.2 para más información), con lo que será necesario elegir un sistema para su implementación. A la hora de analizar cuáles son las necesidades del proyecto y qué tecnología adoptar, se tienen en cuenta principalmente los siguientes aspectos para llevar a cabo una decisión:

- Cantidad y tamaño de datos a almacenar: el sistema tiene que almacenar todos los datos relativos a un curso. Estos en ocasiones son muy numerosos debido a los numerosos indicadores que se almacenan por estudiante y día.
- Flexibilidad y escalabilidad: el sistema debe contemplar la posibilidad de recoger información de varios cursos distintos, existiendo la certeza de que cada uno de los cuales tendrá una estructura completamente distinta (esto es, un distinto número de unidades y subunidades).

Estos dos puntos son fundamentales para tomar la decisión de adoptar un servicio de almacenamiento de datos basado en una base de datos no relacional.

De entre las posibilidades de sistemas de gestión de bases de datos no lineales existentes, se toma la decisión de utilizar MongoDB por los siguientes motivos:

- Es el tipo de base de datos no relacional más usado en la actualidad, con lo que tiene un soporte ya verificado y una comunidad extensa.
- Al estar basado en el almacenamiento de documentos de tipo JSON, es uno de los sistemas más flexibles a la hora del almacenamiento de datos: permite la agrupación de datos de manera estructurada (por ejemplo, Arrays, objetos en formato JSON o estructuras más complejas), y al mismo tiempo es flexible a la hora de almacenar cualquier información no necesariamente estructurada de manera similar a otros documentos de su misma agrupación (colección).
- Facilidad de escalado si hiciera falta en algún momento dado.

2.4.2 Lenguajes de desarrollo y librerías

Los lenguajes de programación utilizados en edX-LIS, Python y R, son lenguajes que siguen siendo ampliamente utilizados en *Data Science*. Sin embargo, su funcionalidad en el uso de la herramienta en ocasiones era complicado debido precisamente a la necesidad de alternar ambos lenguajes en su uso para ir ejecutando la funcionalidad de sus módulos. Por ello, para la creación de una plataforma edX-LIMS que aúne y amplíe la funcionalidad ya existente, se considera necesario abandonar el uso de R y migrar el código y las operaciones que tenían lugar en ese lenguaje a Python. Esta decisión viene dada por la versatilidad y posibilidades que ofrece Python a la hora de construir una plataforma que cumpla las funciones de aplicación *web* y mantenga las funcionalidades ya existentes.

Dadas las circunstancias y los objetivos, se considera la posibilidad de utilizar un *framework* que permita la construcción de una aplicación *web*. Tras analizar las posibilidades existentes, se decide utilizar **Dash** [11] para la construcción de la plataforma. Dash es un *framework* construido a partir del más conocido Flask [16] destinado a la

visualización de datos a través de *Dashboards*. Para ello, integra de forma nativa las funcionalidades de la librería de visualización y generación de gráficos *Plotly* [27] (proyecto hermano de Dash) y la biblioteca de JavaScript *React* [32] para el renderizado de la interfaz de usuario.

La utilización de Dash genera varias ventajas de desarrollo. Aparte de cumplir el objetivo ya mencionado de integración completa del sistema con Python, se trata de una tecnología *Open Source* multiplataforma, y en la cual no es estrictamente necesario dominar las tecnologías de *front-end* para la generación de la interfaz, ya que es algo que se define dentro de la propia definición de elementos, que se realiza en Python. Por el contrario, es precisamente esa estructura, en la que la página de HTML es definida mediante la introducción de elementos en una lista de Python, la que hace algo compleja una maquetación de múltiples elementos.

Para la realización de las múltiples funcionalidades de la plataforma, se utilizan numerosas librerías para determinados procesos:

- *pymongo* [29] como método de conexión entre Python y MongoDB
- *pandas* [26] para la utilización de *Dataframes* (estructuras de almacenamiento de datos de múltiples tipos de carácter bidimensional) en el procesado de datos.
- *flask_login* [17] para la gestión de sesión de usuario.
- *dash_bootstrap* [11] para la maquetación en base a Bootstrap
- *urllib* [39] para el procesado de URLs.
- *cryptography* [10] para la creación de identificadores codificados para cada Estudiante (véase apartado 3.2.3 para más información).

2.4.1 Impacto y satisfacción

El estudio del impacto y la influencia de los servicios que presta la herramienta tuvo lugar a través del análisis y procesado de los datos existentes de la primera edición del curso “Introducción al desarrollo de aplicaciones web” (también llamado “*WebApp*”) ofertado por UAMx [14]. Este curso, actualmente disponible en formato *open* (sin restricciones temporales), tenía una duración planteada de siete semanas en su primera edición. El proceso de acompañamiento se inició al poco de comenzar el curso, recibiendo todos los estudiantes *feedback* sobre su desempeño en el curso vía *email* (progresivamente el envío se restringió a los usuarios con actividad reciente en el curso). El porqué de la elección de este curso se debe a que la tutora del presente TFG es también miembro del equipo docente del susodicho curso en edX, y esto permitía tener acceso a toda la información necesaria para la puesta en funcionamiento de la herramienta. Por ello mismo, los datos de este curso han sido los datos que se han utilizado a la hora de crear y comprobar el funcionamiento de la plataforma edX-LIMS.

Para evaluar cual ha sido el impacto e influencia que este proceso hubiera podido tener en el estudiante, se llevó a cabo una encuesta de satisfacción dirigida a todos los estudiantes del curso. Esta encuesta aportó información en detalle acerca de las consideraciones de los usuarios sobre cuáles han sido los puntos más interesantes de la información que les ha sido proporcionada, y cuál sería la información que les hubiera gustado recibir [7]. Estos datos han sido tomados en cuenta de cara a la creación de edX-LIMS.

3 Diseño

En el siguiente apartado se detalla el sistema propuesto de la plataforma edX-LIMS y los servicios, módulos y diseños que lo conforman.

3.1 Descripción del sistema

La plataforma edX-LIMS (acrónimo de *Learning Intervention Monitoring Service for edX MOOCs*) es una aplicación web que permite realizar un proceso de seguimiento y evolución del aprendizaje de los estudiantes de un curso MOOC de edX. Sus servicios integran una serie de utilidades para los estudiantes del curso online que les facilitan conocer cuál ha sido su evolución y desempeño en el curso a lo largo del tiempo, y se les ofrece una estrategia de acompañamiento en su aprendizaje en relación con las observaciones tomadas mediante la información que devuelve el sistema. Estas utilidades cubren los siguientes procesos:

- Importación y procesado de logs de un curso.
- Generación de indicadores de uso, interacción y evaluación de datos de los estudiantes en relación con su desempeño en un curso.
- Visualización de datos de los estudiantes en formato *Dashboard*.
- Visualización de datos totales de un curso en formato *Dashboard*.
- Visualización de datos del uso de la propia plataforma por parte de los estudiantes.
- Servicio de envío de correos personalizados a los usuarios del curso.

Para llevar a cabo estos procesos y operaciones, se plantea la plataforma con una estructura dividida en servicios en función de su funcionalidad y objetivo de acuerdo con el esquema presentado en la Figura 3-1:

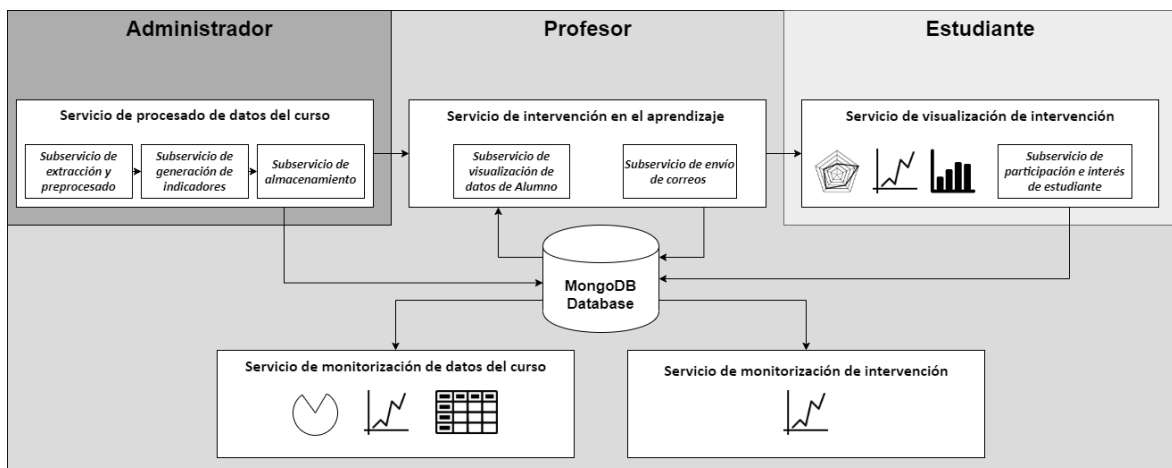


Figura 3-1: Servicios de la plataforma edX-LIMS

La subdivisión visible a través de los tonos de color separados corresponde a la accesibilidad a cada uno de estos servicios por parte de los distintos roles de usuario que se contemplan en la plataforma. Estos roles son los siguientes:

- **Estudiante:** un usuario del curso edX matriculado en el itinerario verificado, es decir, que ha realizado el pago de la matrícula del curso (en adelante estudiante verificado), que ha seguido su desarrollo y, por tanto, ha interactuado con todos los

contenidos del curso (videos, páginas, ejercicios, etcétera). Su acceso se limita a la observación de sus propios datos.

- **Profesor:** un profesor de un curso edX que gestiona toda la información relativa a éste (contenidos, unidades, puntuaciones de ejercicios, etcétera), y por tanto tiene acceso a todos los datos relativos a un curso.
- **Administrador:** un gestor de todos los elementos de la plataforma (cursos, subida de datos, etcétera).

Cada uno de estos roles de usuario tiene acceso a su vez de las funcionalidades propias de roles jerárquicamente inferiores: así, el Administrador tiene pleno acceso a la información disponible para el Profesor y el Estudiante, y el Profesor tiene la del Estudiante.

Una aproximación más en detalle de cada uno de estos servicios se dará en los siguientes subapartados.

3.1.1 Servicio de procesamiento de datos del curso

Este servicio realiza el proceso de ingesta de datos. Su funcionalidad tiene como objetivo llevar a cabo la extracción, limpieza, y procesamiento inicial de datos de estudiantes de un curso.

Subservicio de extracción y preprocesado

La información de un curso de edX viene dada por una serie de ficheros, los cuales serán necesarios para llevar a cabo la importación de datos:

- Fichero de perfiles de usuarios: contienen la información relativa a los datos de cada usuario que estos completan en su perfil de edX. Resulta necesario para poder obtener una relación entre un usuario dado y los logs del curso, y poder identificar de esta manera a quien pertenecen.
- Fichero de certificaciones de usuarios: presenta la información sobre las certificaciones del curso (esto es, la culminación del curso) llevadas a cabo por los usuarios y la fecha en la que tuvieron lugar.
- Logs de curso: cada uno de estos logs se presenta como un fichero de texto de frecuencia diaria (esto es, cada fichero conteniendo todos los logs de un día) con un log por cada línea del fichero. Este log se presenta en formato JSON y representa una acción realizada por un usuario dado y registrada por edX.

A la hora de extraer la información de un curso, y debido a la limitación de información que edX devuelve acerca de usuarios no verificados en el curso de edX (usuarios que no han pagado matrícula del curso, que acceden a todos los contenidos de manera gratuita, excepto a los problemas de evaluación, a los cuales sólo tienen acceso los matriculados en el itinerario verificado), se lleva a cabo el procesamiento de todos los datos filtrados en base a si se trata de usuarios verificados o no, pasando a tener en cuenta únicamente la información de los usuarios verificados.

Subservicio de generación de indicadores

A partir de la información filtrada por el subservicio anterior, este subservicio calcula los indicadores de actividad del curso.

En función de los eventos tenidos en cuenta (véase el anexo B para más información), se identifican y calculan una serie de indicadores de actividad por cada usuario y día. Estos indicadores se muestran en la tabla 3-1.

Nombre	Identificador	Descripción
Number of Events	<i>num_events</i>	Número total de eventos registrados (un evento corresponde a un log)
Number of Sessions	<i>num_sessions</i>	Número de sesiones realizadas en el curso
Video Time	<i>video_time</i>	Tiempo de uso de recursos de vídeo
Problem Time	<i>problem_time</i>	Tiempo de uso de recursos de problemas
Navigation Time	<i>nav_time</i>	Tiempo de uso de recursos de navegación
Forum Time	<i>forum_time</i>	Tiempo de uso de recursos de foro
Total Time	<i>total_time</i>	Tiempo total de uso
Forum Events	<i>forum_events</i>	Número de eventos registrados de actividades de foro
Navigation Events	<i>nav_events</i>	Número de eventos registrados de actividades de navegación
Problem Events	<i>problem_events</i>	Número de eventos registrados de actividades de problema
Video Events	<i>video_events</i>	Número de eventos registrados de actividades de vídeo
Consecutive Inactivity Days	<i>consecutive_inactivity_days</i>	Días consecutivos sin actividad en la plataforma edX
Connected Days	<i>connected_days</i>	Días consecutivos de conexión en la plataforma edX
Different Videos	<i>different_videos</i>	Videos diferentes vistos
Different Problems	<i>different_problems</i>	Problemas diferentes realizados

Tabla 3-1: Tabla de indicadores de edX-LIMS

Adicionalmente, a partir de los logs que devuelven el resultado de la realización de problemas, se identifica cual es la situación actual del estudiante en relación con su desempeño en los problemas evaluables del curso, y con esa información y el conocimiento del peso de cada problema que devuelven esos mismos logs, se realiza el cálculo de la puntuación actual del estudiante.

Subservicio de almacenamiento

Estos indicadores se almacenan en una base de datos no lineal MongoDB para su futura presentación visual.

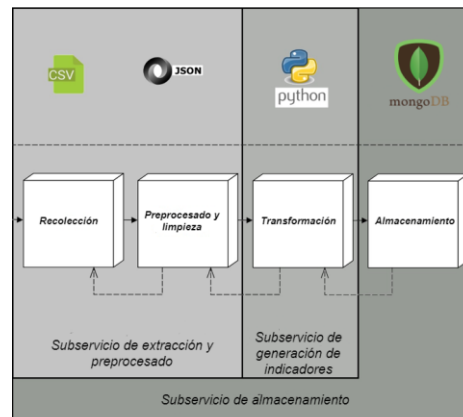


Figura 3-2: Definición del Servicio de procesamiento de datos del curso

3.1.2 Servicio de intervención en el aprendizaje

Este servicio tiene como objetivo presentar los datos recopilados en el servicio de procesamiento de datos del curso. Se persigue un doble objetivo: por un lado, presentar a cada estudiante su propia información para que sea consciente de cuál ha sido su evolución a lo largo del curso, y por otro, ofrecerle esa misma información al Profesor del curso para su análisis y, en caso de ser necesario, dotarle de los medios necesarios para poder generar una estrategia de intervención personalizada para cada estudiante.

Subservicio de visualización de datos de Estudiante

Este subservicio crea una visualización por estudiante que refleja su rendimiento en el curso. Con este objetivo, genera y organiza componentes visuales (gráficas) estructuradas en formato de *Dashboard* con su información.

Subservicio de envío de correos

El objetivo de este subservicio es la gestión del envío de notificaciones a los usuarios mediante correo electrónico. Con ello, se puede llevar a cabo envíos cuyo fin es dar a conocer a los estudiantes la información necesaria para que puedan acceder a sus paneles de aprendizaje o *Dashboard* y avisarles de actualizaciones en los datos existentes para que puedan acceder a sus datos actualizados. Proporciona una serie de herramientas para gestionar el proceso de envío:

- Plantillas predefinidas para disponer del contenido de los correos electrónicos y editables desde la base de datos, así como la posibilidad de edición específica de

cada envío. Integra varios *merge tags* (elementos de nomenclatura específica que se insertan en emails para poder personalizar envíos con datos específicos del usuario destinatario) para poder personalizar el contenido del envío a cada estudiante.

- Elección de destinatarios en base a múltiples filtros basados en actividad, listas de inclusión o exclusión de estudiantes o calificación.
- Visualización del historial de envíos realizados a lo largo del tiempo, sus destinatarios y contenido y los parámetros utilizados para escoger a los usuarios receptores.

3.1.3 Servicio de visualización de intervención

Este servicio proporciona la visualización gráfica de cada *Dashboard* del sistema, pudiéndose visualizar mediante un navegador web. Cada panel de aprendizaje o *Dashboard* muestra el compromiso y el rendimiento del estudiante en el curso en las siguientes visualizaciones:

- La calificación actual del estudiante en el curso, expresado como un valor entre uno como máximo y cero como mínimo. Este se compara con la calificación promedio de todos los estudiantes en el curso. Además, el estudiante puede observar la diferencia entre su calificación y la calificación promedio de todos los estudiantes (véase figura C-1).
- Datos sobre las interacciones del estudiante mostrando sus indicadores diarios. El estudiante puede seleccionar los indicadores para mostrar en los gráficos (véase tabla 3-2). Estos indicadores se pueden comparar con el promedio de estas métricas de todos los estudiantes, tanto en su actividad diaria a lo largo del tiempo como su acumulado en el tiempo. Se dispone de un componente de ayuda para obtener más información sobre la información que representan esos datos (véase figura C-2).
- Datos sobre el rendimiento del estudiante. Se presenta un gráfico de radar que muestra el desempeño del estudiante en las unidades del curso de contenido y un gráfico de barras con su progreso en las tareas calificadas. Además, el estudiante puede comparar sus métricas con los promedios de todos los estudiantes (véase figura C-3).

Subservicio de participación e interés de estudiante

El Servicio de visualización de intervención descrito previamente integra una funcionalidad que registra todas las interacciones de los estudiantes con sus recursos (gráficos, indicadores, selectores y filtros) mediante la creación de eventos que son almacenados en la base de datos MongoDB de edX-LIMS. Esta información será visualizada en el Servicio de monitorización de intervención (apartado 3.1.5).

3.1.4 Servicio de monitorización de datos del curso

Este servicio tiene como objetivo la representación general de los datos totales de un curso de cara a su visualización y análisis por parte del Profesor de este. A través de un *Dashboard* (véase figura C-11), presenta los datos del curso en varios gráficos y tablas de la siguiente forma:

- Datos generales del curso mediante las siguientes métricas: número total de estudiantes (los estudiantes que están inscritos en el itinerario verificado del curso) y estudiantes certificados (los estudiantes verificados que han aprobado el curso porque su calificación es igual o mayor que 0.5), calificación promedio de los estudiantes y calificación promedio de los estudiantes certificados (ya que suele haber grandes diferencias provocadas por la existencia de numerosos usuarios inactivos).
- Datos demográficos de los estudiantes en el curso. Se muestra la distribución de los estudiantes por país, género, edad y nivel académico.
- Datos sobre las interacciones de los estudiantes en forma de sus indicadores calculados diariamente. Es posible aplicar varias operaciones de agregación a estos valores en los gráficos, como el valor promedio o el valor total de las métricas. Se muestran dos gráficos en esta parte, el gráfico de la izquierda muestra las métricas seleccionadas y el gráfico de la derecha muestra estos valores acumulados en el tiempo.
- Datos relacionados con la actividad más reciente de los estudiantes en una tabla y sus calificaciones por unidad y en el curso total. La herramienta de filtrado por determinados valores existente en el Subservicio de envío de correos se encuentra también disponible aquí. Además, es posible acceder al *Dashboard* de cada estudiante desde la tabla.

3.1.5 Servicio de monitorización de intervención

Este servicio permite monitorizar el interés de los estudiantes en la estrategia de intervención en su aprendizaje. Es la representación visual de los datos obtenidos en el Subservicio de participación e interés de estudiante. Muestra las interacciones de los estudiantes en los diferentes gráficos de sus respectivos *Dashboard*. Es posible monitorizar esta información individualmente para un único estudiante o colectivamente para todos los estudiantes o en función de los grupos de usuarios seleccionados para un envío concreto (véase figura 3-11).

3.2 Diseño del sistema

Por utilidad a la hora de ir definiendo la plataforma, se procederá a explicar los distintos modelos de datos y componentes del sistema en un orden jerárquico de interacción *back-front* para permitir la comprensión del diseño del sistema desde su estructura de datos a su capa de aplicación.

3.2.1 Servicio de procesado de datos del curso

Subservicio de extracción y preprocesado

Para el proceso de extracción de datos y generación de indicadores se ha construido la clase **LogProcessor**. Esta clase recibe un fichero zip proveniente de la subida realizada por un usuario con los datos pertenecientes a un curso y realiza la extracción de su contenido (datos de usuarios, certificados y logs) y su procesamiento.

LogProcessor
zip : ZipFile course_data : Object users : Dataframe certs : Dataframe files : [String] log_files : [File]
checkValidity(response, eventType_flag) getDayEvents(file_name) processUsers(users_df, course_id, db) processIndicators(indicators_df, course_id, db) processSeenProblems(seen_problems_df, course_id, db) processSeenVideos(seen_videos_df, course_id, db) processCourseStructure(course_struct_df, course_id, db) getCourseStructure(course_id, db) getProblemData(course_id, db) processEventTrack(event_track_df, course_id, db) getIndicatorsAggr(db, course_id) exportIndicatorsAggr(db, course_id, route) getSeenVideosAggr(db, course_id) exportSeenVideos(db, course_id, route) getSeenProblemsAggr(db, course_id) getSeenProblems(db, course_id) exportSeenProblems(db, course_id, route) getEventTrackAggr(db) getFinalIndicators(db, course_id) exportAllIndicators(db, course_id, route) importCsvToDb(db, db_collection, course_id, csv_route) setCourseInfo(db, course_name, course_id) importLastExistingData(course_id) exportCurrentInactivityDays(db, course_id, route) processUserGradeAvgs(grades_df, course_id, date, db) processUserGrades(grades_df, course_id, date, db) processProblemData(problem_df, course_id, db) processNumProblemsPerStudentPerUnit(nproblems_df, date, db) processNumProblemsPerUnit(nproblems_df, course_id, db) processProblemsToBeDone(problems_df, course_id, date, db) processUndoneAttainableGrade(problems_df, course_id, date, db) processUndoneAttainableGradeAvgs(grades_df, course_id, date, db) processSuccessViability(viability_df, course_id, date, db) processFiles()

Figura 3-3: Definición de la clase LogProcessor

La función principal que realiza el proceso de extracción y procesado, *processFiles*, realiza toda la operativa de conversión y creado de los indicadores y elementos del sistema. Por la naturaleza de estos procesos, que en muchas ocasiones requieren de interacciones con la base de datos, se requiere una serie de funciones auxiliares cuyo único objetivo es esa operatividad, afectando esto tanto a operaciones de naturaleza *get* para poder realizar los cálculos respecto a indicadores acumulativos a lo largo del tiempo, como a las operaciones de tipo *set*, que entrarían dentro de las funcionalidades del Subservicio de almacenamiento de información. En todo caso, ambas operaciones se valen de la funcionalidad que aporta otra clase que será explicada en el siguiente apartado para llevar a cabo su finalidad.

Subservicio de almacenamiento de información

El modelo de datos del sistema toma como referencia la agrupación de datos que se llevaba a cabo en la herramienta edX-LIS y la define en MongoDB, además de añadir la información necesaria para dar soporte a todos los servicios de la plataforma.

Nombre	Datos que contiene
app_users	nombre, password, curso, rol
course_info	nombre, id
course_structure	curso, nombre bloque, problema, peso, intentos máximos
course_users	id, nombre usuario, nombre real, email, idioma, año de nacimiento, género, nivel educativo, metas, modo de curso, estado verificación, cohorte, ciudad, país, fecha creación, último login, nota certificado, estado certificado,

	fecha certificado
email_sendings	usuario plataforma, curso, destinatarios, template, filtros, usuarios específicos seleccionados, asunto, contenido, fecha
email_templates	nombre, bloques
final_indicators	curso, usuario, fecha, num_events, num_sessions, video_time, problem_time, nav_time, forum_time, total_time, forum_events, nav_events, problem_events, video_events, consecutive_inactivity_days, connected_days, different_videos, different_problems
num_problems_student_unit	curso, usuario, fecha, número problemas / unidad, problemas totales
num_problems_unit	curso, unidad, número problemas
problems_to_be_done	curso, usuario, fecha, nombre bloque, problema, peso
seen_problems	curso, usuario, fecha, problema, bloque, puntuación, peso, número intentos
seen_videos	curso, usuario, fecha, video, código
success_viability	curso, usuario, fecha, nota actual, nota máxima, viabilidad
undone_attainable_grade	curso, usuario, fecha, problemas pendientes / unidad
undone_attainable_grade_avgs	curso, usuario, fecha, nota pendiente / unidad
user_filters	usuario plataforma, curso, nombre, filtro
user_grade_avgs	curso, usuario, fecha, nota / unidad
user_grades	curso, usuario, fecha, nota / subunidad
user_tracking	curso, usuario, fecha, nombre, parámetros

Tabla 3-2: Tabla de colecciones de edX-LIMS

La colección *app_users* contiene la información relativa a los usuarios registrados dentro de la plataforma (Profesores y Administradores). Las referencias al dato "usuario plataforma" hacen referencia a usuarios de esta colección.

La colección *course_info* contiene la información relativa al curso: nombre e identificador.

La colección *course_structure* contiene la información relativa a la estructura del curso.

La colección *course_users* contiene información detallada acerca de todos los usuarios del curso sobre los cuales se busca recoger datos (esto es, usuarios verificados). Las referencias de la Tabla 3-2 a "usuario" hacen referencia a los elementos de esta colección.

La colección *email_sendings* es un registro de los envíos realizados mediante el servicio de envío de correos. Contiene todo lo relativo al mismo, incluidos filtros aplicados (en un formato definido en la colección *user_filters*), listas de usuarios a tener en cuenta y contenido del envío.

La colección *email_templates* guarda las definiciones de los *templates* predefinidos del sistema.

La colección *final_indicators* guarda cuales son los indicadores medidos para un usuario en un día concreto. Los parámetros contemplados en su información en el cuadro 3.2 son los definidos en el cuadro 3.1.

La colección *num_problems_student_unit* guarda cuántos problemas pertenecientes a una unidad ha realizado el estudiante. Debido a que la estructura de cada curso de edX puede variar, los parámetros de cada elemento de esta colección vendrán dados por esa estructura, con lo que pueden no ser similares en todos los casos.

La colección *num_problems_unit* guarda el número de problemas disponibles para realizar que tiene una unidad de un curso completo.

La colección *problems_to_be_done* guarda los problemas que todavía no ha realizado un estudiante.

La colección *seen_problems* guarda el registro de cuál ha sido el resultado de un problema que ha sido intentado por un estudiante.

La colección *seen_videos* guarda registro de los vídeos que ha visualizado un estudiante junto con el tipo de visualización que ha realizado.

La colección *success_viability* almacena el cálculo de la nota máxima que un estudiante puede alcanzar como máximo teniendo en cuenta los intentos y resultados de los problemas que ya ha intentado.

La colección *undone_attainable_grade* hace referencia a ese valor máximo alcanzable definido en la colección dividido por subunidad del curso

La colección *undone_attainable_grade_avgs* guarda la misma información que la colección anterior, pero con una nota por unidad

La colección *user_filters* almacena los filtros de estudiantes que un usuario Profesor o Administrador define. El filtro en sí es un array de objetos en el que cada cual es una definición de qué parámetro, operador y valor ha definido.

La colección *user_grade_avgs* contiene la nota de un usuario por cada unidad de un curso dado en una fecha dada.

La colección *user_grades* es el mismo planteamiento que *user_grade_avgs*, pero en una división por cada subunidad.

La colección *user_tracking* almacena la información que se obtiene mediante el Subservicio de participación e interés de estudiante, esto es, los registros de la interacción de un estudiante con su *Dashboard*. En caso de ser un evento de acceso a la página (evento tipo "*pageview*"), el registro no contendrá más datos que el usuario y la fecha en la que ha tenido lugar, pero si interactúa con algún elemento de su *Dashboard*, el valor de los

parámetros almacenará qué tipo de interacción ha sido, los valores que ha introducido y si ha querido realizar una comparación de los valores propios junto con las medias del curso.

En la figura 3-4 se presenta el modelo de datos de la plataforma y las relaciones entre las distintas colecciones que la componen. La nomenclatura utilizada para explicar los tipos de datos más complejos es la siguiente:

- Block (colección *user_templates*). Objeto con la siguiente estructura:
 - *title*: String que define el título del bloque
 - *type*: String que define el tipo del bloque
 - *content*: String que contiene el texto del bloque
- Filter (colecciones *email_sendings* y *user_filters*). Objeto con la siguiente estructura:
 - *fname*: String parámetro a filtrar
 - *fop*: String operador a aplicar
 - *fvalue*: String valor del filtro
- Log (colección *user_tracking*). Objeto cuya estructura es dependiente del evento que lo haya generado:
 - El log “*pageview*” no tiene datos en este campo.
 - El log “*radar-bar*” contiene *date* como un String con la fecha que se haya decidido comparar en el selector de fechas.
 - El log “*indicators*” guarda el campo *indicators* como un array de String con los indicadores seleccionados por el usuario que lo haya generado.

En caso de que exista, siempre guarda el Bool *show_avgs*, que guarda si la interacción con los selectores ha sido comparándose con la media del curso.

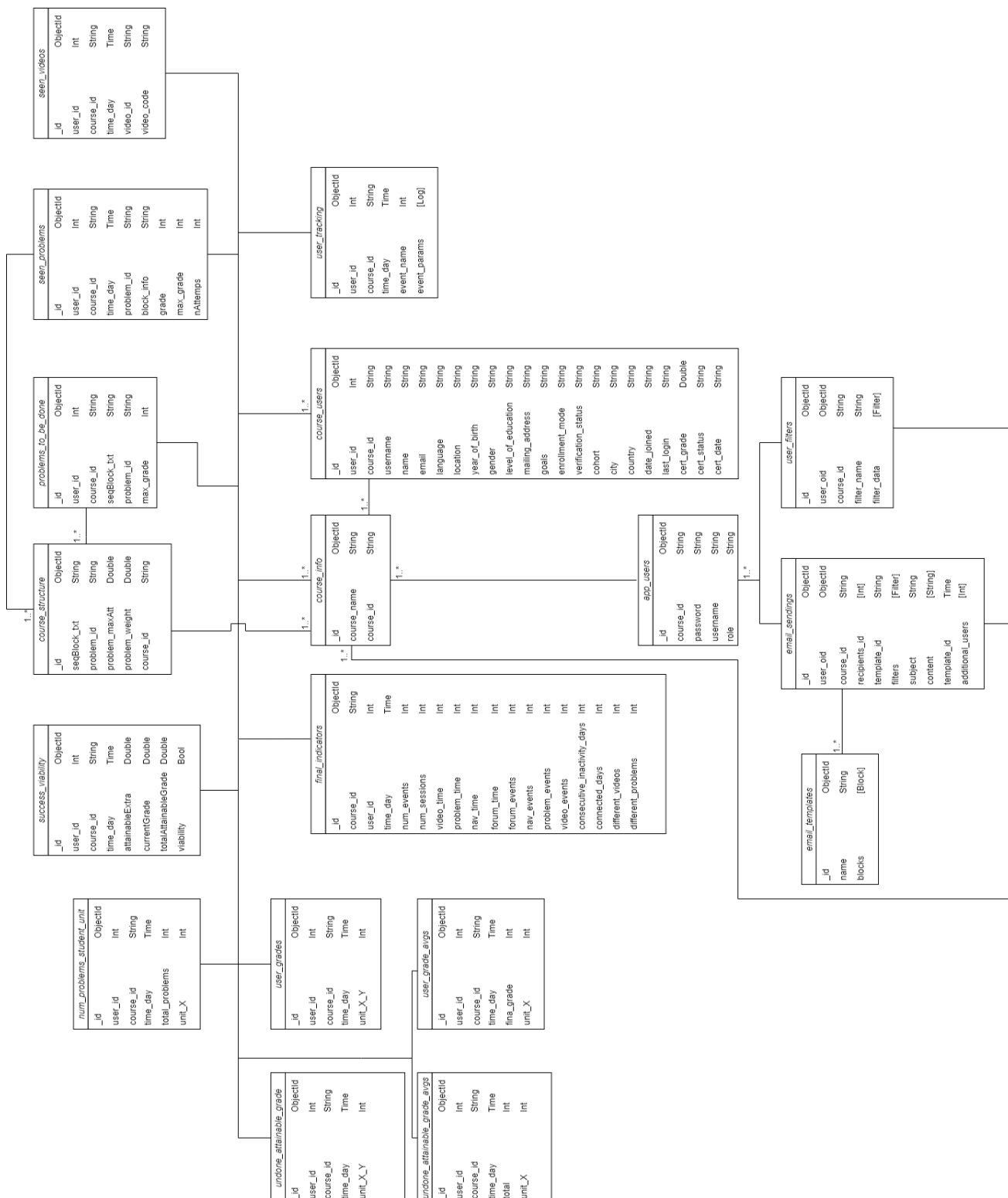


Figura 3-4: Modelo de datos de la plataforma edX-LIMS

De igual forma, en varias colecciones se ha utilizado la nomenclatura *_X* o *_X_Y* para representar a aquellos parámetros cuya existencia es dependiente de la estructura del curso: en estos casos, X representaría el valor numérico de la unidad, e Y el valor numérico de la subunidad de un curso dado (por ejemplo, *unit_5_3* haría referencia a la subunidad 3 de la unidad 5). Un mismo curso tendrá siempre los mismos campos en todos sus documentos, pero diferirán de los de otro curso con una estructura de unidades distinta.

Para llevar a cabo cualquier operación de cualquier naturaleza relacionada con la base de datos, se define la clase **EdxMongoStore**. Esta clase ejerce de pasarela entre los servicios y la base de datos, y se encarga tanto del almacenamiento de datos como de la gestión de consultas para devolver los datos en el formato correspondiente deseado:

EdxMongoStore
db : Database(MongoClient)
<pre> createUser(course_id, username, password) getUserByName(username) updateCourseStructure(course_id, problems) getCourseStructureData(course_id) getProblemData(course_id) saveFinalIndicators(indicators) getDailyIndicators(course_id, measure) saveSeenVideos(indicators) saveSeenProblems(indicators) saveEventTrack(indicators) getCourseInfo(course_id) getCourseStats(course_id) getCourseUsers(course_id) getCourseUsersAvg(course_id, date) updateUsers(course_id, users) getFinalIndicatorsAggrData(course_id) getSeenVideosData(course_id) getSeenProblemsData(course_id) getSeenVideosAggrData(course_id) getSeenProblemsAggrData(course_id) getEventTrackAggrData(course_id) getFinalIndicatorsData(course_id) updateConnectedDaysTime(course_id) checkIndicatorActivity(indicator) getUserInfo(user_id, course_id) getUserLastSeenInfo(user_id, course_id) getUserLastActivityInfo(user_id, course_id) getAnalyticsData(course_id, users_id, avgs) getUserFinalIndicators(user_id, course_id) getUserAvgInfo(user_id, course_id, date) getCourseAvgInfo(course_id, date) getUnitProblemsData(course_id) getUserUnitProblemsData(user_id, course_id, date) getAvgUnitProblemsData(course_id, date) getUserDates(user_id, course_id, element_info) importCollection(info, course_id, collection_name) insertCourseInfo(course_name, course_id) getCurrentInactivityDays(course_id) saveUserGradeAvgs(course_id, time_day, indicators) saveUserGrades(course_id, time_day, indicators) saveProblemData(course_id, problems) saveNumProblemsPerStudentPerUnit(course_id, time_day, users) saveNumProblemsPerUnit(course_id, num_problems) saveProblemsToBeDone(course_id, time_day, problems) saveUndoneAttainableGrade(course_id, time_day, grades) saveUndoneAttainableGradeAvgs(course_id, time_day, grade_avgs) saveSuccessViability(course_id, time_day, success_info) saveUserTracking(user_id, course_id, event_name, event_params) getMailTemplate(template_id) saveUserFilter(user_oid, fname, fdata, course_id, additional_users, cuo) getUserFilters(user_oid, course_id) getFilter(filter_id) saveEmailSettings(user_oid, course_id, template_id, settings_name, filters, mc) saveEmailSending(user_oid, course_id, recipients_id, template_id, filters, additional_users, mc) getEmailSendings(course_id) getEmailSending(sending_oid) getUserEmailSettings(user_oid, course_id) getEmailSetting(setting_id) closeConection() </pre>

Figura 3-5: Definición de la clase EdxMongoStore

La construcción de la clase tiene lugar mediante la instanciación de la conexión con la base de datos a partir de las funcionalidades de la librería *pymongo*, creando un *MongoClient* que queda guardado como un parámetro de la propia clase y sirve como punto de acceso a la MongoDB.

3.2.2 Servicio de intervención en el aprendizaje

Este módulo (y los siguientes a partir de este punto) tiene lugar a través de la interacción con el *frontend* de la plataforma. En la tecnología utilizada para este cometido (Dash) se sigue un patrón MVC (*modelo vista controlador*) para construir la aplicación.

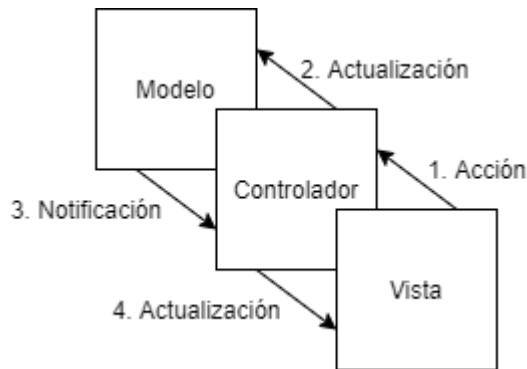


Figura 3-6: Esquema del Modelo Vista Controlador

Por las particularidades de la tecnología utilizada, el sistema se encuentra estructurado en varios archivos, en los que cada uno de ellos contiene tanto una vista (*layout*) como su correspondiente controlador. Con ello, la estructura de navegación de la aplicación consta de un enrutador en su entrada (archivo *app.py* de la figura 3-7) que recoge las peticiones de URL enviadas y las distribuye a sus correspondientes vistas, que gestionarán la información a representar.

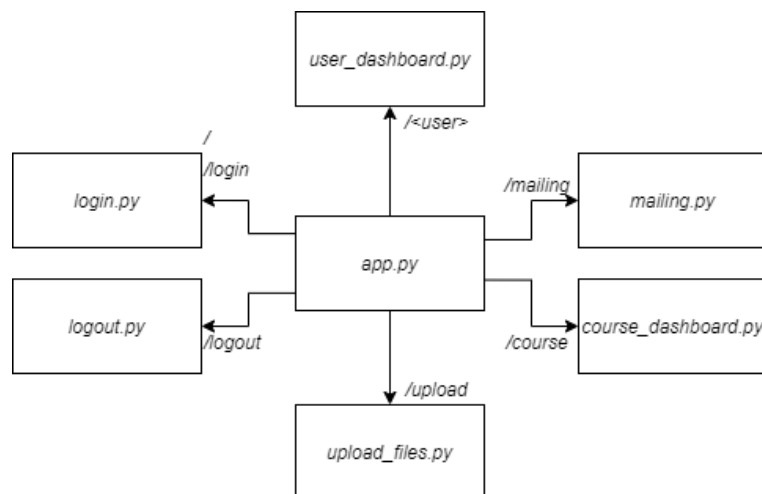


Figura 3-7: Estructura de navegación de edX-LIMS

Subservicio de visualización de datos de Estudiante

Cada uno de los *Dashboard* está formado por una serie de gráficas y selectores que conforman la visualización de los datos. Para ello, Dash integra funcionalidades de representación de gráficas mediante integración nativa con Plotly. En el caso concreto del *Dashboard* de Estudiante, esta representación de datos tiene lugar de la siguiente forma:

- Gráfico indicador (*gauge chart*) con la nota actual del usuario (explicado en más detalle en 3.1.3). La barra azul representa la nota del usuario dentro de una representación en la que la nota por debajo de 0.5 se muestra gris y por encima se muestra verde. La barra amarilla representa la nota media del curso. En caso de que la nota del usuario estuviera por debajo de la media del curso (esto es, que la barra azul no alcanzara la barra amarilla), el indicador delta y el texto explicativo pasarían a ser de color rojo. Un detalle del componente se muestra en la figura C-1.

- Gráficos de líneas (*line chart*) con los datos diarios de los indicadores del estudiante desde el momento en que se detecta que son usuarios verificados, tanto diariamente como en acumulado total (explicado en más detalle en 3.1.3). El Estudiante puede seleccionar los indicadores que quiere visualizar. En caso de seleccionar el *checkbox* del selector (seleccionado en la imagen), se mostrarán también los datos correspondientes a la media del curso en los indicadores seleccionados. Se habilita también un servicio de ayuda si se pincha en el icono con la interrogación. Un detalle del componente se muestra en la figura C-2.
- Gráfico de radar (*radar chart*) con la nota actual por unidad del curso del usuario y gráfico de barras (*bar chart*) con la cantidad de ejercicios realizados y pendientes por unidad (explicado en más detalle en 3.1.3). El Estudiante puede seleccionar la fecha que desee para visualizar la evolución de su progreso entonces (estas fechas se corresponden con las subidas de nuevos datos a la plataforma). En caso de haber seleccionado el *checkbox* (seleccionado en la imagen), se mostrarán también los datos correspondientes a la media del curso en la fecha seleccionada. Un detalle del componente se muestra en la figura C-3

Subservicio de envío de correos

El módulo de gestión de envío de correos se encuentra disponible a través de la URL */mailing*, que lleva a cabo la carga de la vista y los controladores existentes en *mailing.py*.

<i>mailing.py</i>	
Métodos de modelo	<pre> get_past_sendings(course_id) get_sending(sending_oid) translate_user_mergetags(user_id, content, tr_type) send_mail(recipients, content, test, send_to, serv, port) get_user_settings(user_oid, course_id) get_setting(setting_id) save_email_settings(template_id, settings_name, filter_stored, mc) save_email_sending(recipients_id, template_id, filter_stored, users_stored, mc) print_filter(filter_num, filter_values) print_template(t_id, t_title, active) get_mail_template(template_id) </pre>
Métodos de controlador	<pre> show_sections(btn_ns, btn_ps, old_ns, old_ps, old_ps_content, old_sendings) get_sending_preview(sending_cell, sendings_data, old_esp, old_ss) select_template(t1, t2, t3, t4, tp1, tp2, tp3, tp4) add_remove_show_load_filters(btn_add_click, btn_del_click, btn_load_click, btn_add_prev, btn_del_prev, btn_load_prev, oldaf, oldlf, filter_id, oldfn, oldcul, oldcuo) update_delete_filter_btn(filters) save_load_filter(btn_click, nfilter, old, names, ops, values, additional_users, cuo) update_table(new_page_size) update_table(page_current, page_size, sort_by, apply_filter_btn, fnames, fops, fvals, cul, cuo) get_mail_checkout(btn_click, filter_stored, users_stored, cuo, mc, template_id, sort_by) disable_send_button(btn_click) send_email(btn_click, filter_stored, users_stored, cuo, mc, template_id) </pre>

Figura 3-8: Definición de los métodos de *mailing.py*

En la figura 3-8 se representan los métodos existentes en el Subservicio de envío de correos divididos en dos tipos:

- Los métodos correspondientes a la funcionalidad relativa a la operativa de datos (operaciones CRUD de creación, lectura, actualización y borrado propias de los modelos de datos), correspondientes a las operaciones del modelo situados en el espacio superior del cuadro.

- Los métodos definidos como *callbacks* controladores de la vista, los cuales realizan la conexión con la interfaz gráfica y su interactividad, e invocan a los métodos de funcionalidad descritos anteriormente. Al tratarse de programación reactiva, cada uno de estos métodos controla uno o varios elementos de la vista para detectar cambios en su estado y actuar en consecuencia cuando esto ocurra. Se encuentran en la parte inferior del cuadro.

El diseño de la vista del subservicio se muestra en la figura C-5.

El subservicio se subdivide tres apartados:

- Selección de *template* de envío: al seleccionar una de las cuatro opciones, se desplegarán cuadros de texto con el contenido de la opción predefinida (información almacenada en la colección *email_templates*). Esta información podrá ser modificada por el Profesor si así lo deseara. El detalle de la vista de este elemento se muestra en la figura C-6.
- Selección de destinatarios: permite la aplicación, guardado y cargado de filtros tanto por parámetros como por identificadores de usuario para seleccionar a qué usuarios irá destinado el envío, y su visualización en la tabla de la parte inferior. El detalle de la vista de este elemento se muestra en la figura C-7.
- Confirmación del envío. Previo al envío definitivo, el usuario podrá ver una visualización previa de cuál será el mensaje final que vaya a recibir un destinatario como paso final de confirmación antes de realizar el envío. El detalle de la vista de este elemento se muestra en la figura C-8.

Desde la cabecera de la página, el usuario tiene acceso a una vista que se alterna con la descrita hasta ahora (es decir, se encuentra también alojada en */mailing*) y que muestra el historial de envíos previos y el contenido de cada uno de ellos. Un ejemplo de esta vista se encuentra en la figura C-9.

3.2.3 Servicio de visualización de intervención

La generación del *Dashboard* de Estudiante tiene lugar a través de una petición a la URL */<user>*, que redirige a la vista-controlador *user_dashboard.py*. Dado que el Estudiante no requiere de un usuario de la plataforma para poder acceder a su contenido y al mismo tiempo este debe ser únicamente accesible para el Estudiante, el parámetro de la URL *<user>* hace referencia al identificador del usuario (campo *user_id* de la colección *course_users*) encriptado mediante el protocolo de encriptación simétrica proporcionado por el paquete *Fernet* (AES-CBC con clave de 128-bits). Esta URL única le será proporcionada a cada Estudiante mediante el envío de un correo electrónico.

<i>user_dashboard.py</i>	
Métodos de modelo	<pre> get_user_indicators(user_id) get_user_avg_info(user_id, date) get_course_avg_info(date) get_user_property(prop_id, user_data) get_user_radar_dates(user_id) process_radar_info(data) get_unit_problems_data(course_id) get_user_unit_problems_data(user_id, course_id, date, unit_ids) get_avg_unit_problems_data(course_id, date, unit_ids) get_gauge_figure(user_id) save_trace(user_id, course_id, event_name, event_params) get_main_layout(user) toggle_modal(n1, n2, is_open) </pre>
Métodos de controlador	<pre> toggle_modal_et(n1, n2, is_open) toggle_modal_cp(n1, n2, is_open) make_individual_figure(indicators, show_avgs, pathname, search) make_individual_radar(date_value, show_avgs, pathname, search) make_bar_charts(date_value, show_avgs, pathname, search) </pre>

Figura 3-9: Definición de los métodos de *user_dashboard.py*

El *Dashboard* de Estudiante aúna los gráficos descritos en 3.2.2 (figuras C-1, C-2 y C-3), representándolos como aparece en la figura C-4. En la figura C-4 no se encuentra seleccionado el *checkbox* de comparativas con la media del curso, a diferencia de las figuras C-2 y C-3 donde sí está marcado.

Subservicio de participación e interés de estudiante

Asociados a los controladores de los selectores asociados a las gráficas de las figuras C-2 y C-3 se encuentra integrado el método *save_trace*, que genera una traza con la información del selector que está siendo utilizada siguiendo el modelo descrito en 3.2.1 y la guarda en la colección *user_tracking*.

3.2.4 Servicio de monitorización de datos del curso

El *Dashboard* de Profesor se construye en el archivo *course_dashboard.py*, a cuya vista se accede a través de la URL */course*. Un detalle de este *Dashboard* se muestra en la figura C-10.

<i>course_dashboard.py</i>	
Métodos de modelo	<pre> get_course_analytics(course_id, user_id, avgs) get_analytics_label(label, data) print_filter(filter_num, filter_values) get_sending_users(sending_oid) </pre>
Métodos de controlador	<pre> add_remove_show_load_filters_cd(btn_add_click, btn_del_click, btn_load_click, btn_add_prev, btn_del_prev, btn_load_prev, oldaf, oldlf, filter_id, oldfn) update_delete_filter_btn_cd(filters) save_load_filter_cd(btn_click, nfilter, old, names, ops, values) update_table_cd(page_current, page_size, sort_by, apply_filter_btn, fnames, fops, fvals) update_table(new_page_size): make_analytics_figure(indicators, user_id, sending_oid) make_indicator_figure(indicators, measure) </pre>

Figura 3-10: Vista de *Dashboard* de Profesor

El detalle de las distintas secciones del *Dashboard* permite reconocer los elementos descritos en 3.1.4:

- Datos generales del curso.
- Datos demográficos de los estudiantes en el curso en forma de gráficos circulares (*pie chart*).
- Datos sobre las interacciones de los estudiantes en forma de sus indicadores calculados diariamente. Se muestran de manera similar a la figura C-2 perteneciente al *dashboard* de Usuario, como gráficos de líneas que representan datos diarios como en acumulados totales.
- Datos relacionados con la actividad más reciente de los estudiantes en una tabla y sus calificaciones por unidad y en el curso total. La tabla contiene los mismos datos y funcionalidades de filtrado que la existente y ya descrita en el Subservicio de envío de correos (figura C-7).

3.2.5 Servicio de visualización de intervención

La última sección del *Dashboard* de Profesor (figura C-10) representa una serie de datos que no son relativos a la información del curso en sí misma. Esta sección muestra los resultados de la toma de datos de la interacción del Estudiante con su propio *Dashboard* mediante la información recopilada a través del Subservicio de participación e interés de estudiante descrito en 3.1.3 y 3.2.3.

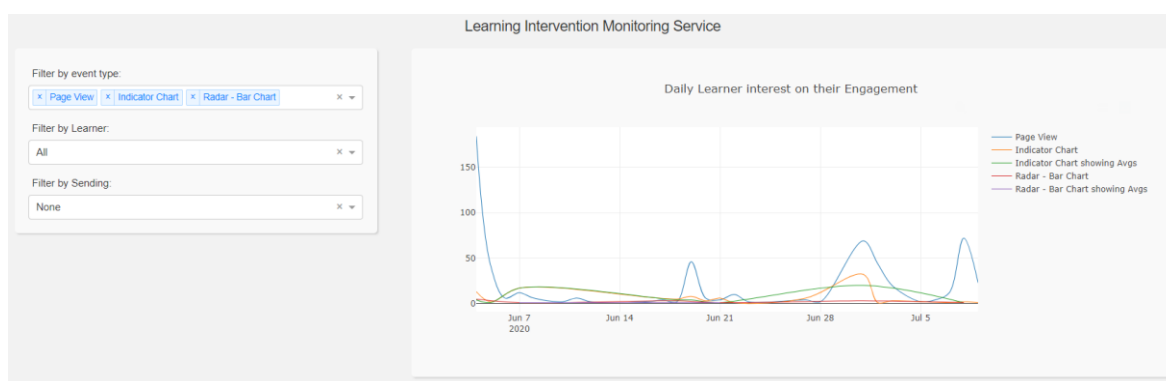


Figura 3-11: Vista del servicio de visualización de intervención

En este caso, los selectores permiten realizar un filtrado de información en relación con el tipo de *log* registrado. Cada uno de ellos representa un tipo de actividad distinto: el acceso al *Dashboard* de Usuario (“Page View”), la interacción con los selectores de la gráfica de la figura C-2 (“Indicator Chart”) o con los de la gráfica de la figura C-3 (“Radar – Bar Chart”). La gráfica automáticamente diferencia y muestra con un registro diferenciado los eventos en los que se haya realizado una interacción que conllevara una comparativa con los datos de las medias del curso. Los selectores adicionales permiten filtrar bien por usuario individual (lo cual permite un hipotético seguimiento de acceso individualizado) o por grupo de usuarios que formen parte de un envío masivo de correos (lo cual permite controlar cual ha sido la respuesta concreta a un envío dado).

4 Desarrollo, implementación y pruebas

Este capítulo desarrolla cual es la estructura del proyecto, su implementación y potencial escalabilidad, y las pruebas a las que sus servicios se han visto sometidas.

4.1 Estructura de ficheros

A la hora de componer la estructura de ficheros del sistema, ha sido necesario cumplir con ciertos requisitos impuestos por la estructura de Dash. Sin dejar de cumplirlos, se ha compuesto una estructura de archivos diferenciada por funcionalidades y objetivos.

La estructura de ficheros del proyecto por carpeta es la siguiente:

- **/:** Raíz del proyecto. Contiene ficheros de configuración y definición de la plataforma.

Fichero	Funcionalidad
<i>app.py</i>	Main del proyecto, enrutador de URLs
<i>config.py</i>	Configuración de la DB y claves de uso del sistema
<i>controls.py</i>	Variables estáticas utilizadas en diversas vistas
<i>server.py</i>	Configuración del servidor
<i>usermodel.py</i>	Modelo de usuario de plataforma para gestión de login y sesión
<i>utils.py</i>	Funciones de utilidad múltiple a lo largo de la herramienta

Tabla 4-1: Estructura de ficheros de /

- **/assets:** Carpeta con los recursos de estilo requeridos por Dash.

Fichero	Funcionalidad
<i>dash-logo.png</i>	Logo de Dash
<i>resizing_script.js</i>	Script de JS para gestionar el redimensionado dinámico de imagen
<i>sl.css</i>	Ficheros de estilo genéricos
<i>styles.css</i>	Ficheros de estilo específicamente creados para edX-LIMS

Tabla 4-2: Estructura de ficheros de /assets

- **/storage:** Carpeta que contiene las clases que gestionan los subservicios de extracción y preprocesado de datos y de almacenamiento.

Fichero	Funcionalidad
<i>edxmongodbstore.py</i>	Definición de la clase EdxMongoStore (subservicio de almacenamiento)
<i>logprocessing.py</i>	Definición de la clase LogProcessor (subservicio de extracción y procesado)

Tabla 4-3: Estructura de ficheros de /storage

4.2 Implementación y escalabilidad

De cara a una ampliación de las funcionalidades existentes, la herramienta admite la posibilidad de escalar sus servicios de varias maneras:

- **Controlar más eventos.** En caso de querer tener en cuenta más tipos de eventos de edX, sería necesario ampliar la funcionalidad de la clase `LogProcessor`. Una vez ahí, el proceso será distinto en función de la funcionalidad que se desee; si se busca añadir nuevos eventos a los cuatro tipos de indicadores que se contemplan actualmente (problemas, videos, foros y navegación), bastaría con incluir el identificador de ese evento dentro de las agrupaciones ya existentes. Si se contemplan otros tipos de modificaciones (por ejemplo, incluir un nuevo grupo de eventos como indicador o subdividir uno existente en dos o más), entonces sería necesario adicionalmente incluir su correspondiente procesado en la propia clase `LogProcessor`, generar sus funciones CRUD en `EdxMongoStore` e incluir sus selectores en los elementos de `controls.py`.
- **Incluir nuevas vistas.** Si se quiere continuar con el modelo planteado y estructurado actual, sería necesario crear un nuevo fichero con la nueva vista dentro de `/views` y referenciarlo en `app.py`, en dónde también sería necesario modificar el enrutador para incluir una nueva URL válida que dirigiera a la nueva vista.
- **Incluir nuevos cursos.** Para realizar este proceso y dado que las rutas actuales parten de la base de que sólo existe un único curso dentro del sistema, sería necesario inicialmente una modificación en el enrutador de las rutas actuales para tener en cuenta el curso en cuestión al que se esté haciendo referencia (por ejemplo, pasando de `/course` a `/course/<id_curso>`, de `/<id_usuario>` a `<id_curso>/<id_usuario>`, y así sucesivamente). A su vez, una posible definición para la creación de un nuevo curso podría pasar por la definición de una ingesta de datos inicial que al menos llevara a cabo la definición del curso para tener en cuenta la división de problemas por unidades. Una vez estuviera definido el identificador y la estructura del curso, se podría llevar a cabo la subida de logs en un proceso similar al ya existente.
- **Añadir funcionalidades nuevas a la vista de *Dashboard*.** Dado que la funcionalidad de la vista se encuentra definida dentro del propio fichero en el que se encuentra la misma, se podría ampliar cualquier funcionalidad ya existente o crear nuevas funcionalidades mediante la codificación de nuevos *callbacks* dentro del fichero. Dependiendo de la funcionalidad, quizás sea necesario también incluir controladores adicionales y/o métodos nuevos de conexión con la base de datos dentro de la clase `EdxMongoStore`.

4.3 Pruebas

Para todo el desarrollo de la aplicación, así como de las pruebas que se detallan a continuación, se han utilizado los datos reales del curso MOOC “Introducción al desarrollo de aplicaciones web” ofertado por la UAM en edX (curso *WebApp*). La realización de estas pruebas ha tenido lugar tanto en un entorno controlado de desarrollo como mediante la observación del comportamiento de usuarios reales del sistema en su despliegue en producción.

4.3.1 Pruebas unitarias

Para controlar el funcionamiento correcto de manera individual de los elementos del sistema, se han tomado en consideración las siguientes pruebas realizadas:

- **Servicio de procesamiento de datos del curso.** La clase LogProcessor es una funcionalidad derivada del proceso existente en edX-LIS sin integración visual existente. Sucesivas pruebas se realizaron en el proceso de integración previo al desarrollo del resto de servicios en los que se llevaba a cabo la importación de datos y su correspondiente almacenamiento. De cara a prevenir errores en la importación de datos, el Subservicio de almacenamiento no procede a ejecutar sus procesos hasta que los procesos anteriores no hayan concluido.
- **Servicio de intervención en el aprendizaje.** Se han realizado pruebas con el fin de comprobar si la generación de información de gráficas se puede ver comprometida con una malformación de los datos o ante valores anómalos de los mismos. De igual forma, se han realizado pruebas de envíos masivos con el fin de comprobar si la funcionalidad del envío en algún momento se altera ante posibles límites en el envío.
- **Servicio de visualización de intervención.** Pruebas de naturaleza similar a las descritas en el servicio anterior: control del comportamiento del *Dashboard* ante datos inválidos o inputs erróneos.
- **Servicio de monitorización de datos del curso.** Control de mostrado de datos anómalos o inputs erróneos en el *Dashboard* de Profesor.
- **Servicio de monitorización de intervención.** Evaluación de la generación de logs circunstancialmente malformados y su impacto en su visualización. Pruebas sobre la generación de logs de manera aleatoria y su correcta visualización en su gráfica correspondiente.

4.3.2 Pruebas de integración y funcionamiento

A la hora de integrar los distintos servicios en una herramienta única, se ha llevado a cabo las siguientes pruebas:

- **Servicio de procesamiento de datos del curso.** De cara a evitar posibles errores, se ha introducido un control sobre el fichero de subida para asegurarse de que existen todos los ficheros que sean necesarios para la correcta creación de datos. El sistema informa al Administrador de que el fichero que intenta subir no se encuentra correctamente formado. Se controla también la duplicación de datos existentes en el sistema.
- **Servicio de intervención en el aprendizaje.** Se controla que el envío de emails tenga un contenido y un destinatario seleccionado para poder enviarlo.
- **Servicio de visualización de intervención.** Se controla que el Estudiante no tenga acceso a un *Dashboard* hasta que exista en el sistema. A su vez, su inclusión en edX-LIMS hace que ya tenga datos correspondientes a la subida que ha realizado, con lo que no se podrá acceder a un *Dashboard* vacío.

- **Servicio de monitorización de datos del curso.** Pruebas de concepto similar a las del servicio inmediatamente anterior al descrito.
- **Servicio de monitorización de intervención.** Observación de la interacción de usuarios reales y su correcto funcionamiento.

4.3.3 Pruebas de sistema

El entorno de desarrollo en el que ha tenido lugar el desarrollo de la aplicación es un ordenador portátil no dedicado con arquitectura de 64 bits, sistema operativo Windows 10, procesador Intel Core I7 de séptima generación y 16GB de RAM. Su despliegue en un entorno de producción ha sido sobre un ordenador de sobremesa no dedicado con arquitectura de 64 bits, sistema operativo Linux Mint 19.2, procesador Intel Core I7 de cuarta generación y 8GB de RAM.

El proceso de prueba en ambos casos ha sido similar: instalación del entorno desde cero (MongoDB, Dash, librerías) y correcta ejecución de todos los servicios del sistema. Su rendimiento es prácticamente similar, únicamente percibiendo una posible y no confirmada pequeña diferencia en el tiempo de procesado de nuevos datos.

5 Validación y resultados

Este capítulo muestra los resultados obtenidos a partir de las experiencias del usuario con la plataforma, así como la valoración de la coordinadora del equipo docente del curso *WebApp* con la misma.

5.1 Valoración de la estrategia de intervención en el aprendizaje

Para evaluar cuál es el impacto que la nueva plataforma ha podido suponer en la percepción del acompañamiento para los estudiantes, se envió a los estudiantes del curso *WebApp* la información correspondiente para el acceso a su propio *Dashboard*, tras lo cual y pasados varios días durante los cuales pudieron familiarizarse con su uso, se les envió a los que habían aprobado (con certificado) una encuesta de satisfacción sobre la plataforma y sus contenidos con el fin de evaluar si han percibido una mejora en su experiencia de aprendizaje online. Los contenidos de la encuesta estaban distribuidos en varias partes a diversas cuestiones acerca de su experiencia con el *Dashboard*.

En primer lugar, es necesario mencionar un factor de importante afectación sobre ella. Durante el periodo en el cual estuvo vigente el Estado de Alarma en España debido a la crisis sanitaria provocada por el COVID-19 (el lapso entre mitad de marzo y finales de junio), el suministro de nuevos datos del curso se interrumpió. Esto hizo que el procesado de estos datos no pudiera llevarse a cabo, y por ende tampoco el envío de emails. Al retomar la actividad una vez finalizado ese tiempo, se dio el caso de que algunos de los estudiantes recibieron información del *Dashboard* por primera vez cuando ya habían terminado el curso. Es por este motivo que las preguntas estén formuladas de tal forma que esos casos indiquen hipotéticamente cómo le podría haber influido.

En segundo lugar, los resultados de la encuesta se mostrarán a continuación mediante diagramas de caja, donde las respuestas dadas por los estudiantes siguen una escala de Likert de 1 a 5, siendo 1 “Totalmente en desacuerdo” con la afirmación planteada, y 5 “Totalmente de acuerdo”.

En una primera parte de la encuesta, se deseaba saber su opinión general sobre la estrategia de acompañamiento recibida, mediante la cual recibirían mensajes que les avisarían que podían ver su desempeño en el curso mediante un *Dashboard*. Los estudiantes valoran en general muy positivamente tener esta oportunidad de ver gráficamente dicho desempeño (véase figura 5-1).

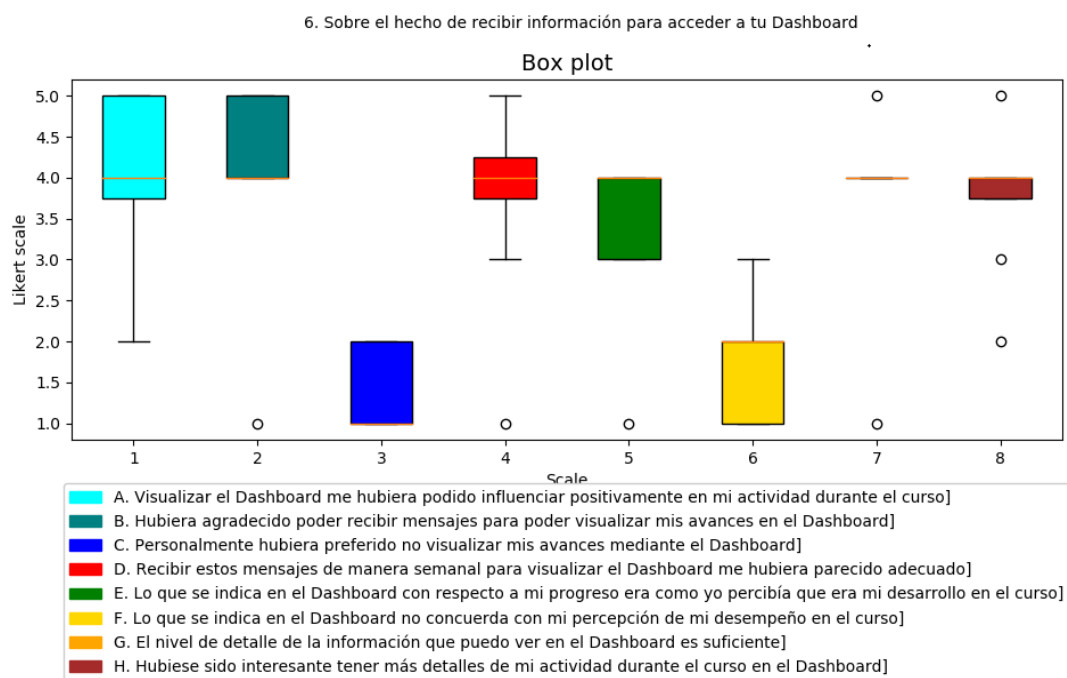


Figura 5-1: Respuesta a preguntas sobre la percepción de los estudiantes sobre su acompañamiento

Los resultados de la encuesta muestran que, por regla general, los estudiantes valoran positivamente su nuevo *Dashboard* y lo toman como una mejora frente a los mensajes de *feedback* que ofrecía la versión de edX-LIS. Se valora especialmente que el *Dashboard* sea interactivo y con la posibilidad de que los usuarios puedan elegir ellos mismos los contenidos que quieran visualizar. Otros puntos bien valorados son los de poder acceder a su evolución temporal, y poder compararse con los datos del resto de estudiantes del curso (véase figura 5-2).

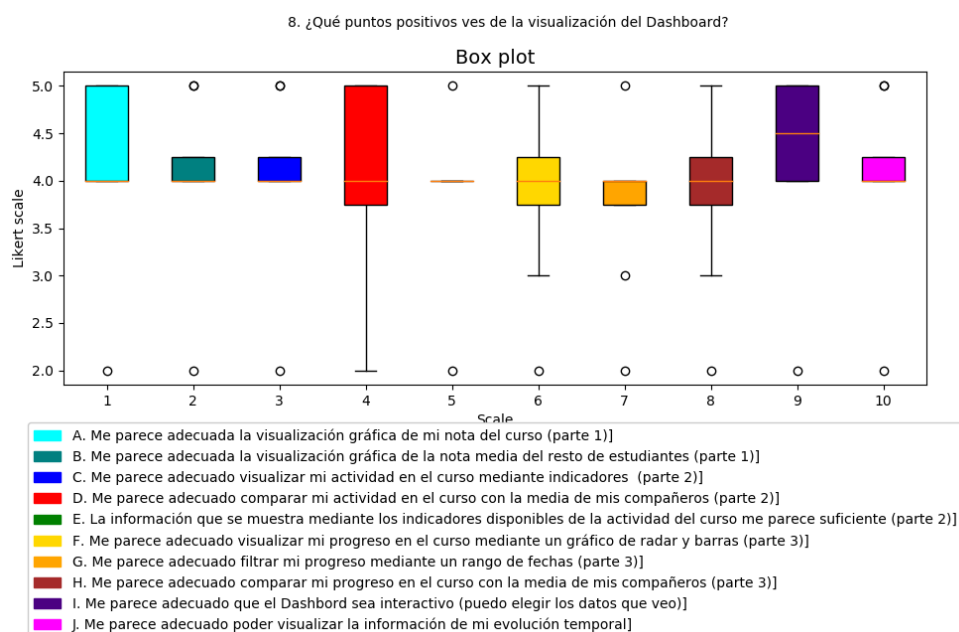


Figura 5-2: Respuesta a preguntas sobre la opinión de los estudiantes acerca de su *Dashboard*

Es importante destacar que se percibe que la existencia del *Dashboard* a lo largo del curso les hubiera ayudado a mantener un ritmo de trabajo constante y a tener en mente que les quedaban actividades y problemas por realizar, así como a sentirse guiados y supervisados en su proceso de aprendizaje, entre otros (véase figura 5-3).

10. El hecho de que hubiera recibido los datos para poder visualizar mi desempeño en el curso mediante el Dashboard antes de que finalizara el curso ...

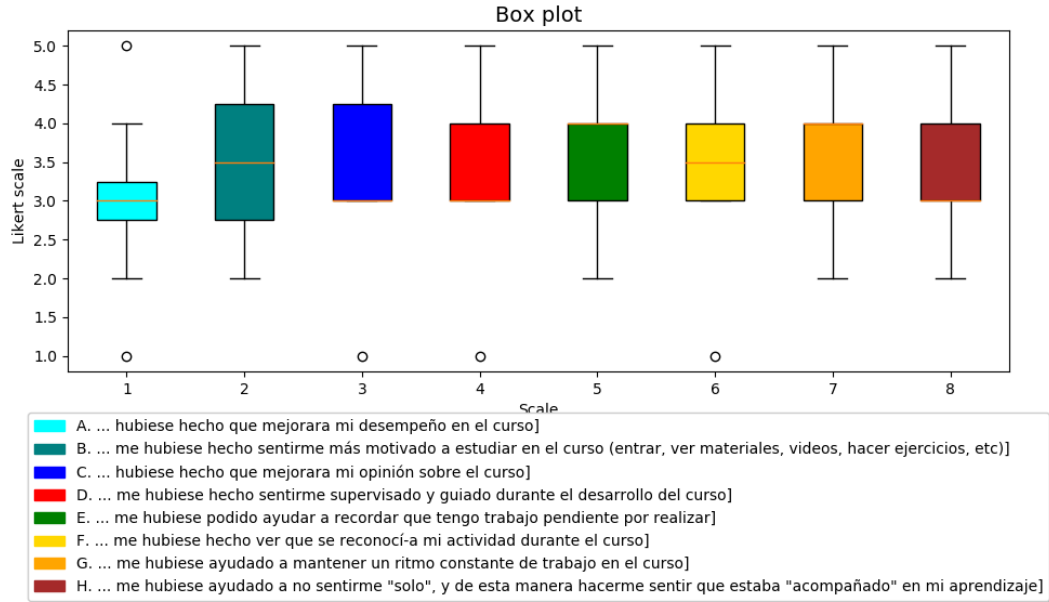


Figura 5-3: Respuesta a preguntas sobre la percepción de los estudiantes acerca de los beneficios de un *Dashboard* de acompañamiento

Finalmente, los estudiantes manifiestan que las visualizaciones que ofrece el *Dashboard* son muy razonables y entendibles en general (véase figura 5-4).

11. Indica con un valor numérico cómo de comprensibles son las visualizaciones que ofrecen las gráficas del Dashboard. Siendo 0: "no se entiende nada y no se cómo utilizar la información que me da", y 10: "se entiende perfectamente y sé cómo utilizar la visualización que me da".

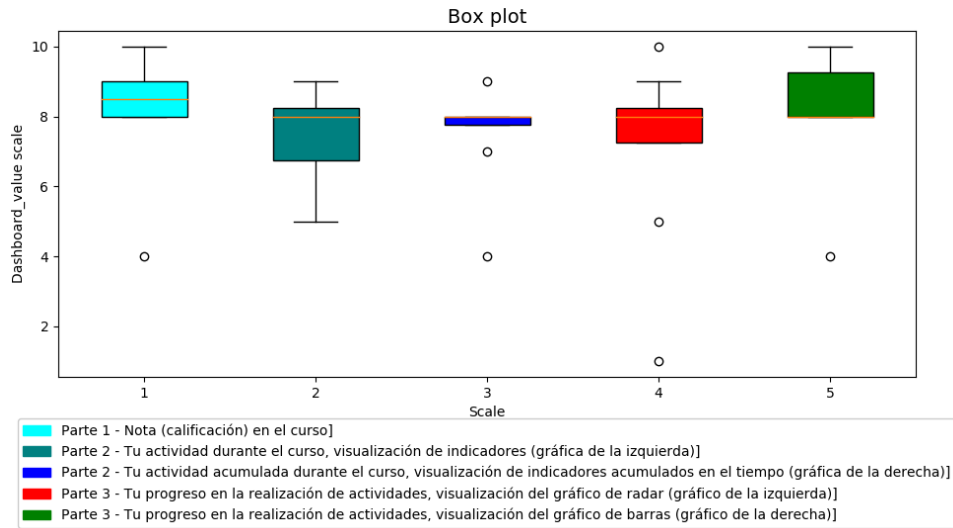


Figura 5-4: Respuesta a preguntas sobre la comprensión de los estudiantes de la información de su *Dashboard* de acompañamiento

5.2 Valoración de la monitorización de la estrategia de intervención

La profesora del curso ha valorado muy positivamente la forma en cómo se muestra la monitorización de los accesos e interacciones de los estudiantes con sus *Dashboard*. Le ha sido muy útil que en el mismo gráfico se muestren los accesos categorizados según las interacciones de los estudiantes a las diferentes partes de su *Dashboard*, así como, que se pueda ver cuántas de estas interacciones conllevan compararse con el desempeño de sus compañeros.

En los primeros envíos se ha podido cuantificar que los estudiantes más motivados a acceder a su *Dashboard* son aquellos que ya han aprobado (con certificado). El mismo día que recibieron su email, accedieron a la plataforma el 25% de los estudiantes aprobados y el 7% de los estudiantes en itinerario verificado y que no han aprobado aún (véase figura 3-11). En opinión de la profesora, son tasas muy aceptables de uso de los *Dashboard*, ya que existe la posibilidad de tomar el envío como un mensaje de *spam* e ignorarlo.

Por último, de acuerdo con la opinión de la profesora, la monitorización y los resultados de la encuesta de las figuras 5-1, 5-2, 5-3 y 5-4 dan razones a continuar realizando este seguimiento, acompañamiento y, en definitiva, estrategia de intervención en el aprendizaje de los estudiantes del MOOC.

6 Conclusiones y trabajo futuro

En este capítulo se presentan las conclusiones y las posibles vías de ampliación de la plataforma a través de hipotéticos trabajos futuros.

6.1 Conclusiones

Los cursos MOOC han sido un método de aprendizaje online de uso creciente durante los últimos años, y por tanto su estudio, enmarcado dentro de las analíticas de aprendizaje (*Learning Analytics*), es una materia cada vez más necesaria e interesante. Dentro de las múltiples plataformas y consorcios existentes, la Universidad Autónoma de Madrid oferta sus cursos desde 2015 en edX, lo que permite la posibilidad de tener acceso a la información generada a partir de la interacción de los estudiantes con los recursos del curso, y por tanto llevar a cabo un análisis de la relación entre curso y estudiante.

Los estudiantes que cursan MOOCs se ven habitualmente afectados por sentimientos negativos que afectan a su desempeño. Estas sensaciones vienen provocadas por diversos hechos o percepciones, como por ejemplo la sensación de que el *feedback* de su actuación en el curso sea escaso o inexistente, o también la falta de interactividad con los docentes del curso. Estas actitudes se reconocen como negativas para el estudiante, y pueden tener como desenlace el abandono del MOOC.

Para mitigar esos factores de riesgo, el sistema propuesto edX-LIMS es una plataforma autónoma que permite el análisis del aprendizaje y el desempeño de los estudiantes de un MOOC, la visualización de dicho análisis y la monitorización del interés de los estudiantes sobre el mismo. Esta plataforma cumple los siguientes objetivos para llevar a cabo esa tarea:

- Importar y procesar logs de un curso.
- Generar indicadores de uso, interacción y evaluación de datos de los estudiantes en relación con su desempeño en un curso.
- Visualizar datos de los estudiantes en formato *Dashboard*.
- Visualizar datos totales de un curso en formato *Dashboard*.
- Visualizar datos del uso de la propia plataforma por parte de los estudiantes.
- Enviar correos personalizados a los usuarios del curso.

Con los servicios que habilitan estos objetivos, se habilitan los siguientes procesos

- La posibilidad de diseñar una estrategia de intervención en los estudiantes y su aprendizaje a través de un sistema que permite la visualización del progreso de todos los estudiantes del curso a través de indicadores y puntuaciones en un *Dashboard* único por usuario.
- El análisis del interés que demuestra un usuario en el servicio que se le proporciona, y por tanto la habilitación de la evaluación del éxito o fracaso de las estrategias diseñadas en el punto anterior. Esto es un indicador adicional que trasciende la relación del usuario con el curso, ya que se monitoriza su relación con los mecanismos de acompañamiento que se le ofrecen.

Esta monitorización es fácilmente accesible por el Profesor en su respectivo *Dashboard*, con el cual podrá identificar de qué forma los usuarios están respondiendo a la estrategia de intervención y evaluar en consecuencia las decisiones a tomar.

A través de una encuesta de satisfacción realizada tras habilitar el acceso a los usuarios a su propio *Dashboard*, los usuarios han respondido positivamente a las visualizaciones de sus datos.

6.2 Trabajo futuro

Las posibilidades de desarrollo de la plataforma son muy amplias dados los criterios de escalabilidad bajo los que se ha creado la plataforma edX-LIMS.

Como evolución natural, la información actual presentada podría ser ampliada con contenidos que aportaran más información a nivel cualitativo. Actualmente se está recogiendo una cantidad de información que, a tenor de las respuestas de la encuesta de satisfacción final realizada (véase apartado 5.1), al usuario le resultaría útil e interesante conocer. Por tanto, un análisis y ampliación de la información existente en el *Dashboard* para incluir estos datos se antoja uno de los primeros pasos dentro de la evolución de edX-LIMS.

Una de las funcionalidades más necesarias ante un eventual crecimiento de la plataforma sería la habilitación de un *Dashboard* para el rol de Administrador. Dadas las limitaciones actuales en relación con el acceso a los datos necesarios, este servicio no se ha planteado como necesario para el funcionamiento óptimo actual. Sin embargo, ante un crecimiento de la plataforma con la inclusión de nuevos cursos, se antoja necesaria una nueva vista que facilite la gestión completa de la plataforma por parte del rol que gestione la información de los cursos.

Otra línea que será necesaria plantear en el momento en el que se pasara a trabajar con múltiples cursos, tiene que ver con encontrar un modelo o servicio que permitiera la comparación entre diversos desempeños. El modelo actual plantea un curso basado en subunidades, pero este planteamiento se trata de un tipo de estructura específica del curso WebApp que no ha sido verificada en otros cursos. En lugar de esperar que existan estructuras similares dentro de otros cursos, sería necesario replantearse si quizás pueda ser necesario un cambio hacia un modelo de datos basado en unidades más estándar y comunes.

Una posibilidad interesante y ambiciosa podría ser la inclusión de las herramientas descritas en el apartado 2.2. Con ello, la plataforma edX-LIMS pasaría a ser un sistema de análisis total y mucho más potente de lo que es en la actualidad. Se estarían cubriendo todos los ámbitos de la analítica (descriptiva, predictiva y prescriptiva), e integraría la visualización de datos del curso, los modelos predictivos que predicen potenciales reacciones negativas de los estudiantes y los métodos de acompañamiento necesarios para evitarlo, todo ello dentro de un sistema autónomo, integrado y que permitiría la interacción y centralización de información entre lo que ahora mismo son servicios independientes. La transformación de estas herramientas, sin embargo, ofrece un gran reto a nivel tecnológico, dado que haría falta transformar cada una de las tecnologías utilizadas para reconstruirlas dentro del *stack* tecnológico de edX-LIMS, lo cual es un proceso que en ocasiones puede ser muy complejo.

Referencias

- [1] Alamán X.; Carro, R. M.; Cobos R. et al. GHIA: Modelado de Estudiantes, Analíticas del Aprendizaje, Atención a la Diversidad y e-Learning. En Revista Iberoamericana de Informática Educativa, Número 30, Julio-Diciembre 2019, pp. 1-12
- [2] Bases de datos no relacionales En línea. Disponible: <https://aws.amazon.com/es/nosql/> [Último acceso: julio 2020]
- [3] By The Numbers: MOOCs in 2019 En línea. Disponible: <https://www.classcentral.com/report/mooc-stats-2019/> [Último acceso: julio 2020]
- [4] Cassandra En línea. Disponible: <https://cassandra.apache.org/> [Último acceso: julio 2020]
- [5] Cobos, R.; Macías, V. edX-MAS: Model Analyzer System. TEEM 2017, Cádiz, 2017.
- [6] Cobos, R.; Olmos, L. (2018). A Learning Analytics Tool for Predictive Modelling of Dropout and Certificate acquisition on MOOCs for Professional Learning. En IEEE International Conference on Industrial Engineering and Engineering Management.
- [7] Cobos, R.; Ruiz García, Juan-C. Improving Learner Engagement in MOOCs using a Learning Intervention System: a research study in Engineering Education. Computer Applications in Engineering Education. ISSN 1099-0542 Special Issue of Distance Learning & Globalization. Aceptado. In press.
- [8] Cobos, R.; Soberón, J. A proposal for Monitoring the Intervention Strategy on the learning of MOOC learners. Learning Analytics Summer Institute Spain 2020, LASI-Spain 2020. (In press CEUR Conference Proceedings).
- [9] CRAN mailr En línea. Disponible: <https://CRAN.R-project.org/package=mailR> [Último acceso: julio 2020]
- [10] cryptography En línea. Disponible: <https://cryptography.io/en/latest/> [Último acceso: julio 2020]
- [11] Dash En línea. Disponible: <https://plotly.com/dash/> [Último acceso: julio 2020]
- [12] Dash Bootstrap Components En línea. Disponible: <https://dash-bootstrap-components.opensource.faculty.ai/> [Último acceso: julio 2020]
- [13] DB-Engines Ranking En línea. Disponible: <https://db-engines.com/en/ranking> [Último acceso: julio 2020]
- [14] edX: Introducción al desarrollo de aplicaciones web En línea. Disponible: <https://www.edx.org/course/introduccion-al-desarrollo-de-aplicaciones-web-2> [Último acceso: julio 2020]
- [15] Elasticsearch En línea. Disponible: <https://www.elastic.co/es/elasticsearch/> [Último acceso: julio 2020]
- [16] Flask En línea. Disponible: <https://flask.palletsprojects.com/en/1.1.x/> [Último acceso: julio 2020]
- [17] Flask Login En línea. Disponible: <https://flask-login.readthedocs.io/en/latest/> [Último acceso: julio 2020]
- [18] LASI 2020 Program En línea. Disponible: <https://lasi20.snola.es/posts/program.html> [Último acceso: julio 2020]
- [19] Macías, V. Herramienta para el modelado predictivo en entornos educativos en línea. Trabajo de Fin de Grado. Escuela Politécnica Superior, Universidad Autónoma de Madrid, Junio 2017.
- [20] Martínez-Monés, A.; Dimitriadis, Y.; Aquila-Natale, E.; Álvarez, A., Caerio-Rodríguez, M.; Cobos, R.; Conde-González, M. A.; García-Peñalvo, F. J.; Hernández-Leo, D.; Menchaca, I.; Muñoz-Merino, P. J.; Ros, S.; Sancho-Vinuesa, T. (2020).

- Achievements and challenges in learning analytics in Spain: The view of SNOLA. RIED. Revista Iberoamericana de Educación a Distancia, 23(2), (preprint). doi: <http://dx.doi.org/10.5944/ried.23.2.26541>
- [21] Microsoft Data Platform En línea. Disponible: <https://www.microsoft.com/en-us/sql-server/> [Último acceso: julio 2020]
 - [22] MongoDB En línea. Disponible: <https://www.mongodb.com/> [Último acceso: julio 2020]
 - [23] MySQL En línea. <https://www.mysql.com/> [Último acceso: julio 2020]
 - [24] Olmos, L. Sistema informático para el análisis de datos en entornos educativos. Trabajo de Fin de Grado. Escuela Politécnica Superior, Universidad Autónoma de Madrid, Febrero 2018.
 - [25] Oracle Database Cloud En línea. Disponible: <https://www.oracle.com/database/> [Último acceso: julio 2020]
 - [26] Pandas En línea. Disponible: <https://pandas.pydata.org/> [Último acceso: julio 2020]
 - [27] Plotly En línea. Disponible: <https://plotly.com> [Último acceso: julio 2020]
 - [28] PostgreSQL En línea. Disponible: <https://www.postgresql.org/> [Último acceso: julio 2020]
 - [29] Pymongo En línea. Disponible: <https://pymongo.readthedocs.io/en/stable/> [Último acceso: julio 2020]
 - [30] Python Imaging Library (PIL) En línea. Disponible: <https://pillow.readthedocs.io/en/stable/index.html> [Último acceso: julio 2020]
 - [31] Qué es una base de datos relacional En línea. Disponible: <https://www.oracle.com/es/database/what-is-a-relational-database/>, [Último acceso: julio 2020].
 - [32] React En línea. Disponible: <https://es.reactjs.org/> [Último acceso: julio 2020]
 - [33] Rodríguez, A. Herramienta informática de apoyo al usuario en entornos educativos en línea. Trabajo de Fin de Grado. Escuela Politécnica Superior, Universidad Autónoma de Madrid, Junio 2018.
 - [34] Romero, C; Ventura, S. Educational data mining and learning analytics: An updated survey. WIREs Data Mining Knowl Discov. 2020;e1355. <https://doi.org/10.1002/widm.1355>
 - [35] SNOLA En línea. Disponible: <https://snola.es/> [Último acceso: julio 2020]
 - [36] SOLAR En línea. Disponible: <https://www.solaresearch.org/> [Último acceso: julio 2020]
 - [37] UAMx. En línea. Disponible: <https://uamx.uam.es/> [Último acceso: julio 2020]
 - [38] UAMx en edX. En línea. Disponible: <https://www.edx.org/school/uamx> [Último acceso: julio 2020]
 - [39] urllib En línea. Disponible: <https://docs.python.org/3/library/urllib.html> [Último acceso: julio 2020]
 - [40] Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, <https://ggplot2.tidyverse.org>.

Anexos

A Encuesta de satisfacción con edX-LIS

Durante los meses de octubre de 2019 a enero de 2020 se realizó una encuesta de satisfacción sobre la herramienta edX-LIS dirigida a los usuarios del curso *WebApp*. Esta encuesta fue respondida por un total de 24 personas.

Las principales conclusiones del resultado de la encuesta fueron las siguientes:

- **Pregunta 6 (“A continuación se harán afirmaciones sobre los mensajes recibidos con asunto ¿Estamos contigo!”):** El envío de información de seguimiento semanal es aplastantemente positiva a ojos de los usuarios. Lo único destacable fuera de esta opinión es la pregunta 6.H; una parte considerable de usuarios no considera interesante información adicional a la ya mostrada.
- **Pregunta 8 (“Si hubieras podido elegir la periodicidad de recibir estos mensajes, ¿cuál crees que es la mejor para ti?”):** Una amplia mayoría prefiere un seguimiento de evolución semanal.
- **Pregunta 10 (“¿Qué puntos positivos ves en cada una [de las versiones de Dashboard de edX-LIS (figura A-1)]?”):** Comparativa entre versiones:
 - 10.A - 10.B (“Los elementos de la Versión 1/2 me parecieron fácilmente interpretables.”): Facilidad de interpretación prácticamente similar (misma cantidad de personas lo ven fácil, una más ve difícil la versión 2 respecto a la 1).
 - 10.C - 10.D (“La información de la Versión 1/2 era suficiente.”): Siendo mayoría los que creen que la información es suficiente en ambas versiones, son más los que lo creen de la versión 2.
 - 10.E (“(Versión 1) Ver mi desempeño comparado con el de los demás estudiantes me ha sido de utilidad para seguir adelante con el curso.”): La información de la versión 1 no ha sido de especial utilidad para seguir el curso (genera indiferencia).
 - 10.H - 10.I (“Los gráficos e imágenes de la versión 1/2 son muy complejos/difíciles de interpretar.”): Ni la versión 1 es compleja de entender ni la versión 2 es incompleta en cuanto a su información.



Figura A-1: Imagen de *Dashboard* de edX-LIS utilizada en la encuesta

- **Pregunta 11** (“Si pudieses elegir los elementos a recibir en estos mensajes, ¿con qué elementos te quedarías?”): La mayoría preferiría poder disponer de la información de ambas versiones al mismo tiempo.
- **Pregunta 12** (“El hecho de recibir estos mensajes de acompañamiento...”): Los mensajes han sido especialmente eficaces para motivar al alumno, amén de hacerle sentir acompañado y mejorar su opinión sobre el curso. Sin embargo, los que perciben que haya influido positivamente en su desempeño son casi tantos como los que no lo han percibido.
- **Pregunta 13** (“¿Qué más añadirías en la información recibida? ¿Qué echas en falta para complementar y saber sobre tu desempeño en el curso? Puedes añadir lo que creas oportuno (puedes indicarnos tipos de gráficos, o qué textos serían adecuados, etc.)”): La mitad de las opiniones consideran correcto el estado actual. La mayor parte de las recomendaciones e ideas no se refieren a métricas, sino a ideas de otro tipo. La única opinión directamente relacionada con gráficos sugiere que "Toda métrica relativa es interesante".
- **Pregunta 13b** (“Estamos pensando en facilitar esta información de acompañamiento a través del acceso a una página Web (*Dashboard* o Panel de control privado). En este caso si se recibiera la información de acompañamiento de esta manera.”). Un *Dashboard* Web estaría muy bien visto. Sin embargo, existen opiniones muy dispares sobre la querencia de seguir recibiendo esta información vía email (aunque quizás estén malentendiendo esta pregunta, puesto que pueden estar entendiendo recibir la información en sí, y no la notificación de la actualización de información en el *Dashboard*, que es lo que se mandaría).
- **Pregunta 14** (“¿Cómo crees que te han ayudado los mensajes de acompañamiento recibidos?”): La mayor ayuda percibida por los usuarios de los emails es el recordatorio de realizar el curso y la sensación de reconocimiento de estar completándolo.

- **Pregunta 15 (“¿Qué nota has obtenido al finalizar el curso? o si no has terminado aún ¿qué nota tienes ahora?”):** Los usuarios que han completado esta pregunta tienen una media aproximada (no real, sino basada en esta pregunta) de 7.3157894736842105, con una moda de 8.
- **Pregunta 17 (“A lo largo del curso, ¿te hubiera interesado saber si, dada tu actividad y progreso, si tus posibilidades de obtener un certificado eran altas (pudiendo adquirirlo más adelante)?”):** A la mayoría les hubiera interesado tener la información de si su posibilidad de obtener certificado era alta.
- **Pregunta 18 (“Por favor, puedes dejarnos un mensaje al equipo docente a continuación en relación con este tema de la encuesta.”):** Muchos de los mensajes son de felicitación por el curso, por lo que se entiende que han quedado satisfechos con su realización.

B Desglose de eventos para la generación de indicadores

A la hora de tomar eventos de los logs del curso para su procesamiento y almacenamiento, se definen cuatro grandes tipologías de evento. Para cada una de estas familias se toman una serie de logs relacionados con ella como indicadores relevantes para ser tenidos en cuenta para los procesos desglosados en los apartados 3.1.1 y 3.2.1. Estas tipologías son las siguientes:

- **Eventos de video.** Relacionados con la actividad realizada por el usuario en cuanto a la visualización de recursos de vídeo del curso. Los eventos de edX registrados como eventos de esta familia se muestran en la Tabla B-1.

Identificador edX	Evento que registra
<i>hide_transcript</i>	Ocultado de subtítulos de video en navegador
<i>edx.video.transcript.hidden</i>	Ocultado de subtítulos de video en aplicación móvil
<i>load_video</i>	Finalización del renderizado de vídeo en navegador
<i>edx.video.loaded</i>	Finalización del renderizado de vídeo en aplicación móvil
<i>pause_video</i>	Pausado de vídeo en navegador
<i>edx.video.paused</i>	Pausado de vídeo en aplicación móvil
<i>play_video</i>	Reproducción de vídeo en navegador
<i>edx.video.played</i>	Reproducción de vídeo en aplicación móvil
<i>seek_video</i>	Selección de un momento del vídeo en navegador
<i>edx.video.position.changed</i>	Selección de un momento del vídeo en aplicación móvil
<i>show_transcript</i>	Mostrado de subtítulos de video en navegador
<i>edx.video.transcript.shown</i>	Mostrado de subtítulos de video en aplicación móvil
<i>speed_change_video</i>	Cambio de la velocidad de reproducción del vídeo
<i>stop_video</i>	Finalización del video en navegador
<i>edx.video.stopped</i>	Finalización del video en aplicación móvil
<i>video_hide_cc_menu</i>	Cambio del idioma de subtítulos
<i>video_show_cc_menu</i>	Selección de subtítulos
<i>hide_transcript</i>	Ocultado de subtítulos de video en navegador

Tabla B-1: Eventos de vídeo de edX tomados en cuenta en edX-LIMS

- **Eventos de problema.** Son eventos relativos a la interactividad del usuario con los problemas de evaluación del curso. Sus eventos relacionados se muestran en la Tabla B-2.

Identificador edX	Evento que registra
<i>edx.problem.hint.demandhint_displayed</i>	Petición de pista para resolver un problema
<i>edx.problem.hint.feedback_displayed</i>	Petición de feedback tras envío de respuesta
<i>problem_check</i>	Comprobación de problema correcto
<i>problem_check_fail</i>	Comprobación de problema erróneo
<i>problem_graded</i>	Problema puntuado correcto
<i>problem_rescore</i>	Problema vuelto a puntuar correcto
<i>problem_rescore_fail</i>	Problema vuelto a puntuar erróneo
<i>problem_reset</i>	Reinicio de respuesta de un problema
<i>problem_save</i>	Petición de guardado de problema
<i>problem_show</i>	Mostrado de respuesta de problema
<i>reset_problem</i>	Reinicio de problema correcto
<i>reset_problem_fail</i>	Reinicio de problema erróneo
<i>save_problem_fail</i>	Guardado de problema erróneo
<i>save_problem_success</i>	Guardado de problema correcto
<i>showanswer</i>	Mostrado de respuesta de problema correcto

Tabla B-2: Eventos de problema de edX tomados en cuenta en edX-LIMS

- **Eventos de navegación.** Son los relativos a la interacción del usuario con los elementos que permiten la navegación entre unidades del curso. Los eventos registrados se muestran en la Tabla B-3.

Identificador edX	Evento que registra
<i>page_close</i>	Cierre de página
<i>seq_goto</i>	Navegación entre unidades no consecutivas
<i>seq_next</i>	Navegación a la siguiente unidad
<i>seq_prev</i>	Navegación a la anterior unidad
<i>edx.ui.lms.link_clicked</i>	Click en link de hipertexto
<i>edx.ui.lms.outline.selected</i>	Selección de subsección
<i>edx.ui.lms.sequence.next_selected</i>	Navegación a la siguiente unidad
<i>edx.ui.lms.sequence.previous_selected</i>	Navegación a la anterior unidad
<i>edx.ui.lms.sequence.tab_selected</i>	Selección de navegación a otra unidad

Tabla B-3: Eventos de navegación de edX tomados en cuenta en edX-LIMS

- **Eventos de foro.** Tienen en cuenta la actividad y participación del usuario en los foros de discusión del curso. Sus respectivos eventos se encuentran en la Tabla B-4.

Identificador edX	Evento que registra
<i>edx.forum.comment.created</i>	Crear comentario
<i>edx.forum.response.created</i>	Crear respuesta
<i>edx.forum.response.voted</i>	Votar respuesta
<i>edx.forum.searched</i>	Buscar hilo
<i>edx.forum.thread.created</i>	Crear hilo

<i>edx.forum.thread.voted</i>	Votar hilo o respuesta
<i>edx.forum.thread.viewed</i>	Ver hilo

Tabla B-4: Eventos de foro de edX tomados en cuenta en edX-LIMS

C Vistas de edX-LIMS

El siguiente Anexo muestra las figuras correspondientes a las vistas de la plataforma edX-LIMS.

Elementos del *Dashboard* de Estudiante

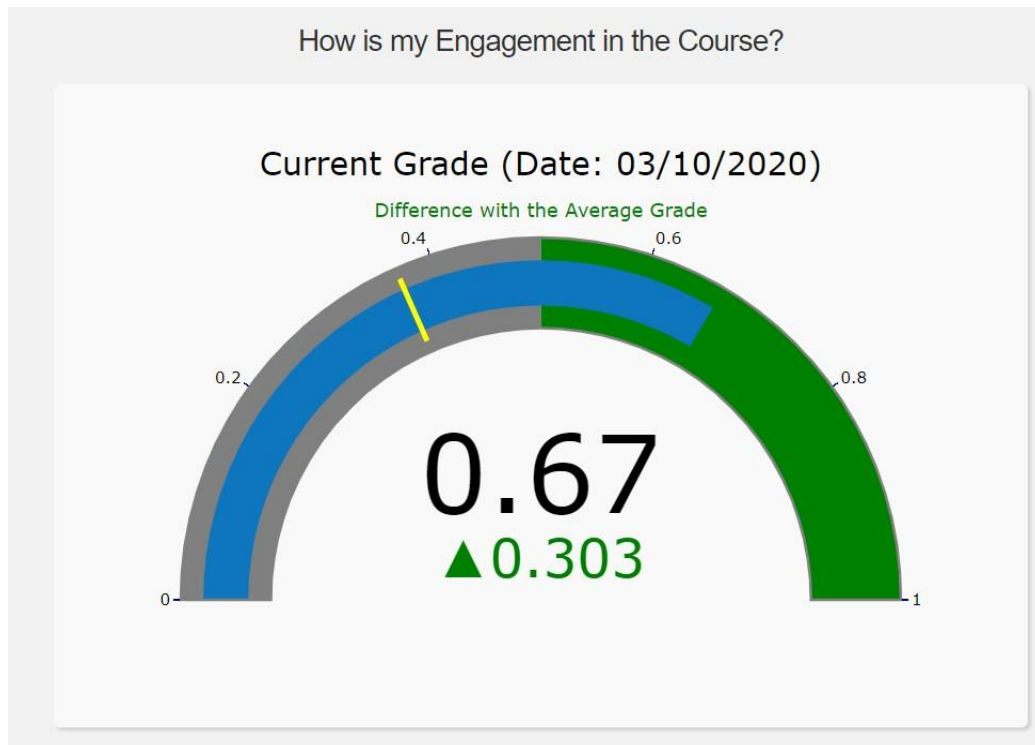


Figura C-1: Gráfico indicador de nota actual de usuario

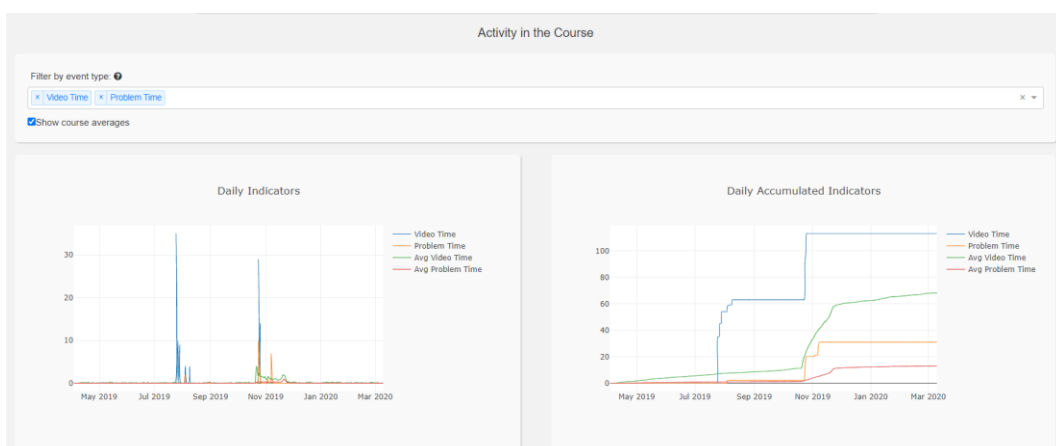


Figura C-2: Gráficos de línea con los valores de los indicadores del usuario

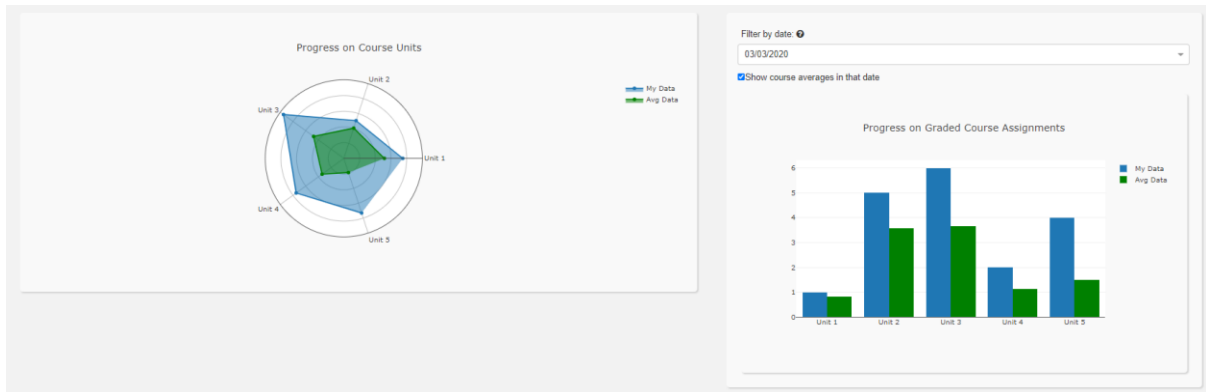


Figura C-3: Gráficos de radar y de barra con el progreso por unidad del Estudiante

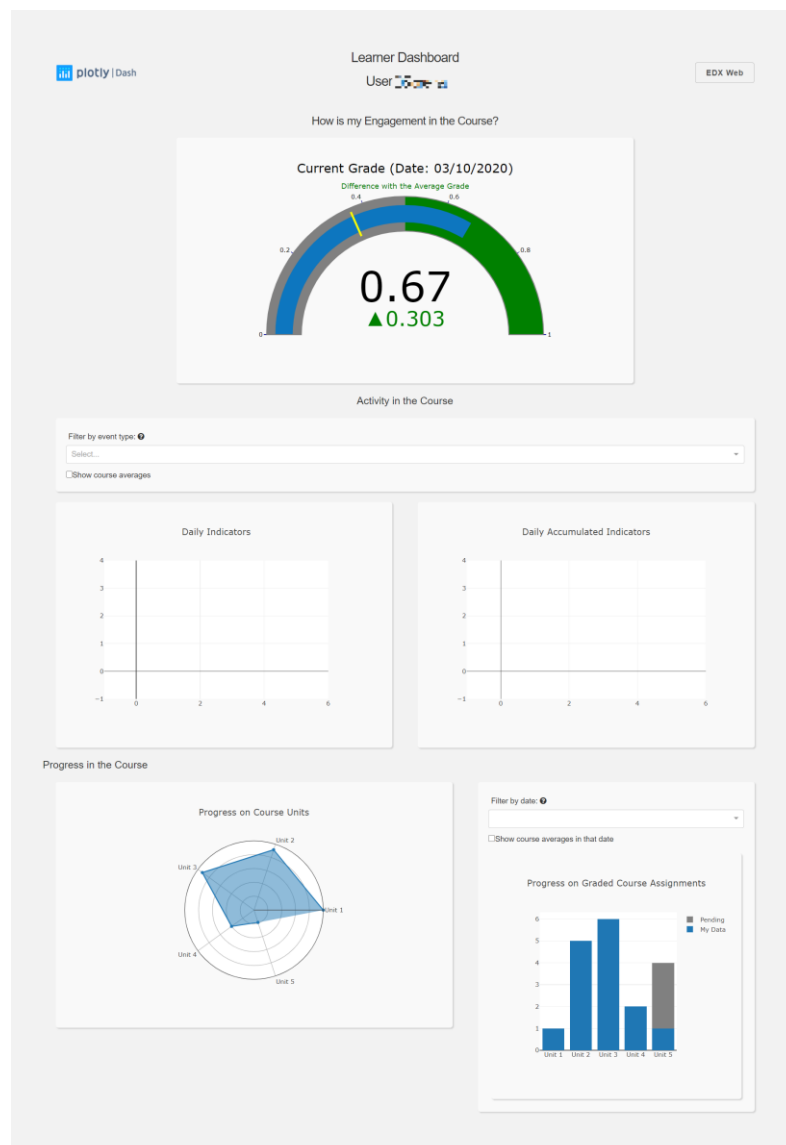


Figura C-4: Vista de Dashboard de Estudiante

Elementos del servicio de *mailing*

plotly | Dash

WebApp Course

Course Dashboard

Mailing

SEND NEW EMAIL

PAST SENDINGS

1. Select your mail template

Option 1

Option 2

Option 3

Option 4

Remember that the following mergetags are available:
[user_name] : User nickname
[user_dashboard_url] : Link to the user Dashboard

2. Select your recipients

Number of results:
10 results

Filters:
User Id
ADD FILTER
DELETE LAST FILTER
SAVE FILTER
Filter name:
Saved filters:
Saved filters:
adad
LOAD FILTER
Additionally, you can paste the user IDs you want (comma separated) and they will be added to or excluded from the sending:
☒ Include ☐ Exclude
APPLY FILTER

User Id	User Name	Certificate Date	Certificate Status	Consecutive Inactivity Days	Unit 1	Unit 2	Unit 3	Unit 4	Unit 5	Final Grade	Learner Dashboard
adad	adad	2019-11-12 19:42:18	downloadable	4	1	0.92	0.92	0.4	0.39	0.67	Access
adad	adad	2019-11-03 16:51:13	downloadable	107	0.5	0.54	0.61	0.75	0.3	0.51	Access
adad	adad	2019-10-26 13:05:42	downloadable	115	0.75	0.51	0.95	0.75	0.73	0.74	Access
adad	adad	2020-01-14 20:36:43	downloadable	43	0.75	0.78	0.9	0.5	0.42	0.69	Access
adad	adad			339	0	0	0	0	0	0	Access
adad	adad	2019-11-15 03:20:09	downloadable	46	0.75	0.82	0.87	0.75	0.25	0.67	Access
adad	adad	2020-01-22 02:20:17	downloadable	48	0.5	0.36	0.71	0.8	0.35	0.51	Access
adad	adad			97	0	0	0	0	0	0	Access
adad	adad	2019-11-20 20:16:26	downloadable	105	1	0.36	0.68	0.75	0.39	0.48	Access
adad	adad			78	0.25	0.28	0.82	0.1	0	0.35	Access

1 / 69

3. Confirmation

Your email will be send to 684 recipients

CHECK EMAIL

Figura C-5: Vista de /mailing

1. Select your mail template

Option 1

Option 2

Option 3

Option 4

Remember that the following mergetags are available:
[user_name] : User nickname
[user_dashboard_url] : Link to the user Dashboard

Subject

¡Estamos contigo!

Block 1

¡Hola [user_name]!

Nos alegró volver a escribirte hace dos semanas.

Como te recordamos en el anterior email, hace unos meses te facilitábamos resultados del seguimiento de tu actividad en el curso semanalmente.

Block 2

Ahora tenemos la forma de facilitarte lo que te mandábamos por email hasta el día que aprobaste el curso a través de un Dashboard interactivo en este enlace: [user_dashboard_url]

Figura C-6: Detalle de selección de *template* de email

Number of results:

10 results

Filters:

User Id

=

Value

ADD FILTER

DELETE LAST FILTER

SAVE FILTER

Filter name:

Saved filters:

adass

LOAD FILTER

Additionally, you can paste the user IDs you want (comma separated) and they will be added to or excluded from the sending:

Include

Exclude

APPLY FILTER

User Id	User Name	Certificate Date	Certificate Status	Consecutive Inactivity Days	Unit 1	Unit 2	Unit 3	Unit 4	Unit 5	Final Grade	Learner dashboard
00 00 00	00	2019-11-12 19:42:10	downloadable	4	1	0.92	0.92	0.4	0.19	0.67	Access
00 00 00	00 00 00	2019-11-09 16:51:13	downloadable	107	0.5	0.54	0.61	0.75	0.3	0.51	Access
00 00 00	00 00	2019-10-26 13:05:42	downloadable	115	0.75	0.51	0.95	0.75	0.73	0.74	Access
00 00 00	00 00 00	2020-01-14 20:36:43	downloadable	43	0.75	0.78	0.9	0.5	0.42	0.69	Access
00 00 00	00 00 00			339	0	0	0	0	0	0	Access
00 00 00	00 00 00	2019-11-15 09:20:09	downloadable	46	0.75	0.82	0.87	0.75	0.25	0.67	Access
00 00 00	00 00 00	2020-01-22 02:20:17	downloadable	48	0.5	0.36	0.71	0.8	0.35	0.51	Access
00 00 00	00 00 00			97	0	0	0	0	0	0	Access
00 00 00	00 00 00	2019-11-26 20:16:26	downloadable	105	1	0.36	0.68	0.75	0.19	0.48	Access
00 00 00	00 00 00			78	0.25	0.28	0.82	0.1	0	0.35	Access

«

1

/

69

»

Figura C-7: Detalle de selección de destinatarios de email

3. Confirmation

Your email will be sent to 684 recipients

CHECK EMAIL

The following email preview is based on the first recipient selected:































































































































To:                                                                                                                                    

Figura C-8: Detalle de confirmación de contenido de email

Sendings history				
	num_recipients	old	subject	time_d
	320	Se091e2422380012c640138	(Estamos contigo)	2020-06-04T18:11:16.737Z
	4	Se091f332380012c640145	(Estamos contigo)	2020-06-04T18:20:03.295Z
	8	Se091f502380012c640148	(Estamos contigo)	2020-06-04T18:20:00.278Z
	94	Se092a02380012c640157	(Estamos contigo)	2020-06-04T18:34:40.899Z
	87	Se0923a22380012c640168	(Estamos contigo)	2020-06-04T18:35:10.388Z
	39	Se0924202380012c640169	(Estamos contigo)	2020-06-04T18:41:12.481Z
	39	Se0924902380012c64016c	(Estamos contigo)	2020-06-04T18:43:07.345Z
	352	Se0933a02380012c640170	(Estamos contigo)	2020-06-04T18:47:39.375Z
	90	Se0937c52380012c64018a	(Estamos contigo)	2020-06-04T18:56:53.045Z
	90	Se0938022380012c64018d	(Estamos contigo)	2020-06-04T18:59:20.373Z
	90	Se0938e02380012c64018f	(Estamos contigo)	2020-06-04T19:01:34.580Z
	60	Se09396f2380012c640190	(Estamos contigo)	2020-06-04T19:03:43.342Z
	60	Se09398f2380012c640193	(Estamos contigo)	2020-06-04T19:04:15.923Z
	69	Se0939dc2380012c640196	(Estamos contigo)	2020-06-04T19:06:14.627Z
	1	Se0939e0c0b02c64019f33a	(Estamos contigo)	2020-06-04T19:12:45.000Z

Figura C-9: Vista de historial de envíos previos

Elementos del *Dashboard* de Profesor

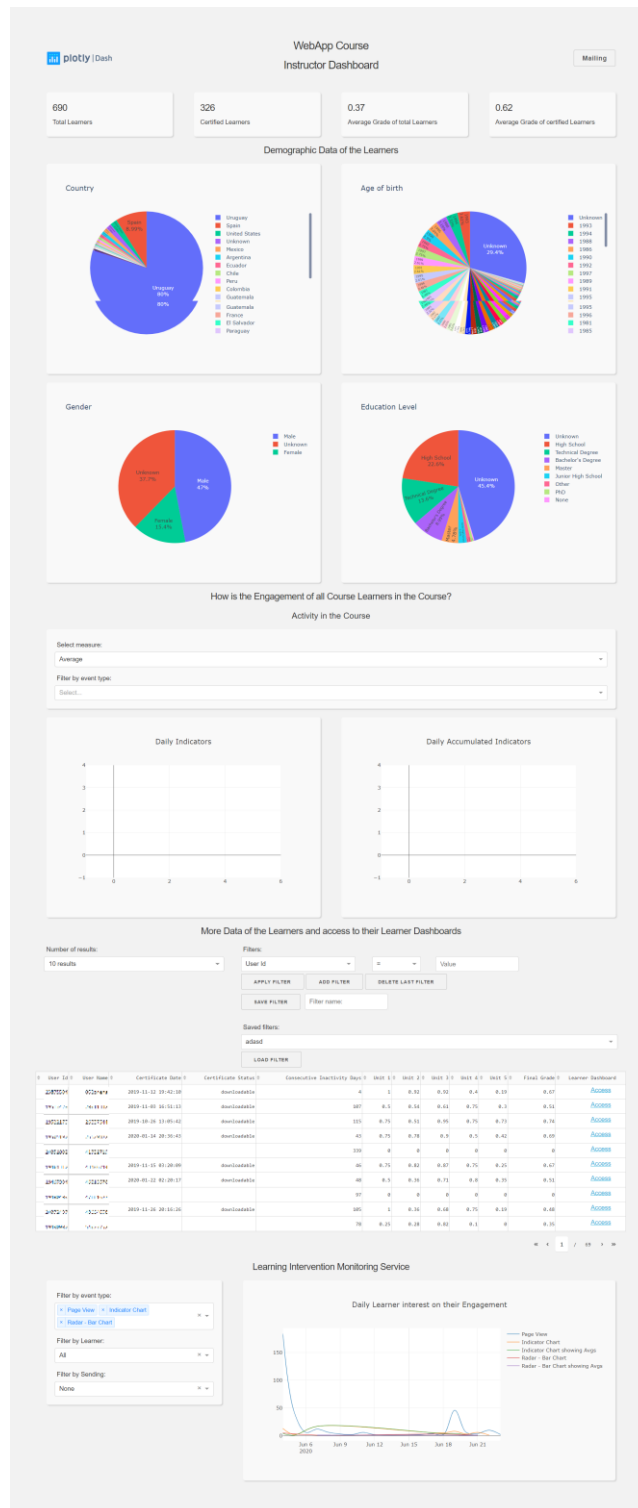


Figura C-10: Vista de *Dashboard* de Profesor

D “A proposal for Monitoring the Intervention Strategy on the learning of MOOC learners”

Cobos, R.; Soberón, J. *A proposal for Monitoring the Intervention Strategy on the learning of MOOC learners. Learning Analytics Summer Institute Spain 2020, LASI-Spain 2020. (In press CEUR Conference Proceedings).*

Este Anexo contiene el *paper* [8], realizado como complemento del presente Trabajo de Fin de Grado, el cual ha sido aceptado y presentado en el congreso LASI 2020 (*Learning Analytics Summer Institute Spain 2020*) [18]. El presente Anexo es autocontenido: se respeta la notación original, con lo que todas las referencias y figuras incluidas en él hacen referencia a su propio contenido, y no al del presente Trabajo de Fin de Grado, el cual ya ha sido referenciado previamente. Del mismo modo, se respeta el idioma original de la publicación (inglés), diferenciándolo del idioma del presente Trabajo de Fin de Grado.

June 16, 2020

12:00-13:00 (CET) - Keynote: Sensing-based Analytics to Support Learning Design (Michalis Giannakos)
Chair: Yannis Dimitriadis
[Link to the session / Enlace a la sesión](#)

13:15-14:30 (CET) - Session Papers 3
Chair: Ainhoa Álvarez Arana
[Link to the session / Enlace a la sesión](#)

Specifying information dashboards' interactive features through meta-model instantiation (Full paper)
Andrea Vázquez-Ingelmo, Francisco José García Peñalvo, Roberto Theron and Alicia García-Holgado

A proposal for Monitoring the Intervention Strategy on the learning of MOOC learners (Full paper)
Ruth Cobos and Juan Soberón

Conceptual framework for process-oriented feedback through Learning Analytics Dashboards (Short paper)
Iñigo Arriaran Olalde and Nagore Ipiña Larrañaga

Figura D-11: Programa de LASI 2020

Abstract

Insufficient feedback and lack of interaction among instructors and learners affect negatively learner retention and engagement in MOOCs. These factors, paired with the feeling of isolation, have a high influence on learners who do not complete the course in which they have enrolled. To overcome this situation, we propose a system to provide periodically MOOC learners with visual information on a Web-based Learner Dashboard, showing them their progress and engagement in the MOOC. This provided information is part of an Intervention Strategy on the learning of these learners. The system offers MOOC instructors to access to a Web-based Instructor Dashboard that shows the interest in this service by the learners and, therefore, facilitates evaluating the success or failure of the Intervention Strategy.

Keywords: Learning Analytics, Massive Open Online Course, Learning Strategy, Web-based Technology, Feedback, Engagement, Intervention.

1 Introduction and motivation

As times goes by, Massive Open Online Courses (MOOCs) have continuously grown in supply (number of courses offered), demand (number of learners enrolled) and relevance in e-Learning discipline [8]. Nowadays, they are accepted as a valid source of educational resources, even sometimes as an official formation recognized by higher education institutions.

In the case of the University Autónoma of Madrid (UAM, Spain), it offers several MOOCs at edX since 2015 (<https://www.edx.org/school/uamx>). In this article, the approach presented was tested in the MOOC: “Introduction to Development of Web Applications” (WebApp MOOC), as in previous works.

Several research studies agree on identifying that insufficient feedback and lack of interaction among instructors and learners affect negatively learner retention and engagement in MOOCs [5][11]. These factors, paired with the feeling of isolation, have a high influence on learners who do not complete the course in which they have enrolled. Therefore, it is necessary to address these issues to adequately direct research to learners at risk of dropout.

It is well known that the main research discipline that can provide ideas to solve the presented problem is Learning Analytics [7][10]. We find a wide variety of Learning Analytics tools in different higher education institutions for prediction and visualisation, among others, in different aspects of learning in MOOCs [6][9]. Furthermore, one of the authors of this article has developed different learning analysis tools previously [1][2][3][4] to solve related problems.

With the aim of decreasing the mentioned feeling of isolation in MOOCs, at UAM we have developed a Learning Analytic system for periodically providing feedback messages to the learners. This system is called edX-LIS (acronym of Learning Intervention System for edX MOOCs). It provides to the course instructors the feedback messages, based on the learner performance in the course and have information on their progresses and recommendations about how they could improve their performance. Therefore, the instructors can send these mentioned messages to the MOOC learners.

In order to test the intervention strategy on learners’ learning supported by edX-LIS a research study conducted with the learners of the first run of the mentioned MOOC. The obtained results from this study have demonstrated that the intervention strategy proposed had a positive impact in the motivation, persistence and engagement of the learners in the MOOC.

Nevertheless, we obtained information about the usefulness of the intervention strategy thanks to the answers obtained from a satisfaction questionnaire, that, only few of the enrolled learners that received these feedback messages answered.

Therefore, to provide monitoring of the intervention strategy on the learning of the MOOC learners, we propose in this article the extension of the previous system to a Web-based Learning Analytics System called edX-LIMS (acronym of System for Learning Intervention and its Monitoring for edX MOOC). edX-LIMS provides, among other

services, MOOC learners with an easy way to visualise their engagement in the course in a “Web-based Learner Dashboard” and it provides MOOC instructors with an easy visualisation of the learners’ interests in the aforementioned visualisation in a “Web-based Instructor Dashboard”.

The structure of this article is as follows: in the next section, we present a detailed description of the proposed Learning Analytics tool: edX-LIMS; in section 3, we present the current use of the tool in a case study with learners of the aforementioned MOOC. Finally, the article ends with conclusions and future works.

2 The approach proposed: a System for Learning Intervention and its Monitoring for edX MOOCs

We propose a Web-based Learning Analytics System called edX-LIMS, which provides an intervention strategy on the learners’ learning and the monitoring of the mentioned strategy by the instructors. We have developed the system for any edX MOOC; however, we have tested it on a specific course (see Section 3 for more details).

In addition, in MOOCs at edX Platform, only the learners enrolled in the verified itinerary (i.e. the learners who pay a fee) can carry out course assessment activities and, therefore, they can be evaluated and get a final grade in the course. Our approach deals with only these verified learners. Therefore, in this article, when we mention MOOC learners, we refer exclusively to MOOC learners enrolled in the verified itinerary.

The user roles for a specific MOOC in our system are: i) Learner (any learner who is enrolled in the verified itinerary), ii) Instructor (any member of the course instructors team) and iii) Admin (any instructor who maintain and manage the course data in the system). Five interconnected services provide the system functionality. These services are coded in Python language (<https://www.python.org/>) and presented in the next subsections.

2.1 Introduction to the system proposed

The system services are (see Fig. 1):

- Course Data Processing Service: the system, based on course data, calculates the learners’ indicators, which represent a summary of the learners’ interactions.
- Learning Intervention Generation Service: the system, based in the learners’ indicators, provides the intervention strategy on the learners’ learning. That is, it generates interactive Web-based Dashboards (one per learner). Therefore, it sends the corresponding information for the access of these Dashboards to the learners via email.
- Intervention Visualisation Service: The learners have the opportunity to access to their Learner Dashboards, which show them their engagement and performance in the course.
- Course Data Monitoring Service: the system allows instructors to monitor summary course data.
- Learning Intervention Monitoring Service: the system allows instructor to monitor the interests of the learners in their received Learners Dashboards.

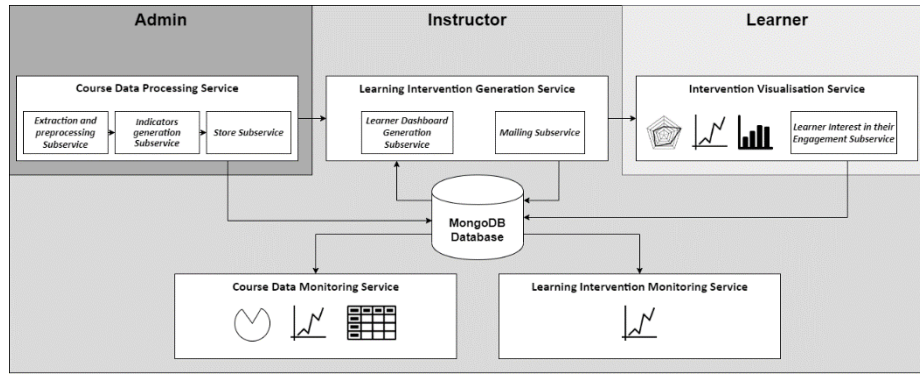


Fig. 1. edX-LIMS services.

The system provides their users three types of Web-based Dashboards. Each one of them can be accessed by the corresponding role user and provides access to the appropriated system services:

- **Learner Dashboard:** Intervention Visualisation Service.
- **Instructor Dashboard:** the service contained in the Learner Dashboard, Learning Intervention Generation Service, Course Data Monitoring Service and Learning Intervention Monitoring Service.
- **Admin Dashboard:** the services contained in the Instructor Dashboards and Course Data Processing Service.

2.2 Couse Data Processing Service

This service allows extracting, cleaning, selecting and preprocessing course data (MOOC log tracks) to obtain the relevant learners' activity. Moreover, it supports the creation of input variables or indicators and data storage management. The following subservices compose this service.

Extraction and Preprocessing Subservice

It is designed to extract and clean course data, preprocess and change its representation. In more detail, in this subservice, we build the learners' activity matrix considering their registered events. We filter events data by type and user. We only collect the events by users who are verified learners.

Indicators Generation Subservice

This service calculates the indicators for the course. These indicators measure for each learner the following: time spent at the course, time invested in graded assignments, videos and forum, number of sessions and events registered, number of days connected, the number of consecutive inactivity days, and the number of different graded assignments solved and videos accessed. In addition, it calculates the current final grade for each learner considering the partial grades of their solved graded assignments.

Store Subservice

The generated indicators are stored as data collections in a MongoDB Database.

2.3 Learning Intervention Generation Service

This service generates the intervention strategy to the learners' learning considering the learners' indicators extracted in the previous service. The following subservices compose this service.

Learner Dashboard Generation Subservice

This subservice creates a visualisation per learner that reflects their performance in the course. With this aim, it generates and organises visual components (graphs) with the use of the framework Dash (<https://plotly.com/dash/>), which is based on Flask and React.js.

Mailing Subservice

The aim of this subservice is to send the information to access to their Learner Dashboards to the learners via email. In addition, this subservice provides the use of templates to configure the content of the emails, and the recipients, among others.

2.4 Intervention Visualisation Service

This service provides the graphical visualisation of each Learner Dashboard, which can be visualised in any web browser. Each Learner Dashboard shows the engagement and performance of the learner in the course in the following visualisations:

- The current grade of the learner in the course (see Fig. 2), which is a value between one to zero (one is the maximum value and zero the minimum one). This value (number in black shown as a blue line in the graph) is compared with the average grade of all the learners in the course (yellow line in the graph). In addition, the learner can see the difference between their grade and the average grade of all the learners (number in green).

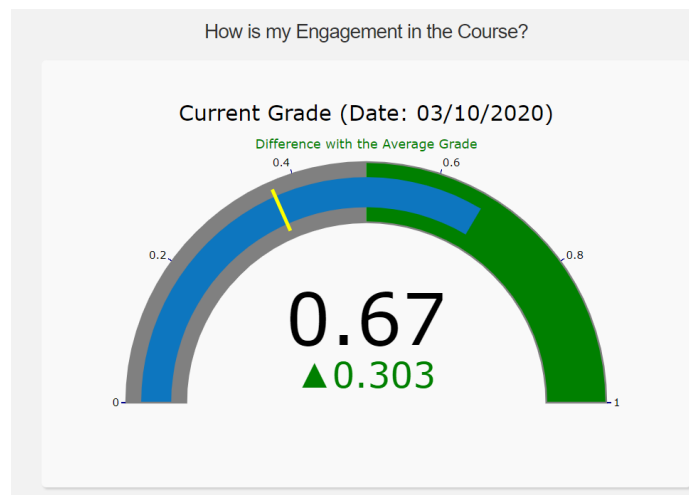


Fig. 2. Graph with the current grade of the learner.

- Data about the learner interactions in form of their calculated indicators daily (see Fig. 3). The learner can select the indicators to show in the graphs. These indicators can be compared with the average of these metrics of all the learners. The left graph shows the selected metrics and the right graph shows these values accumulated in

time. The learner can click in the question mark icon to access help, i.e. a pop-up window with descriptions of the values in these graphs.

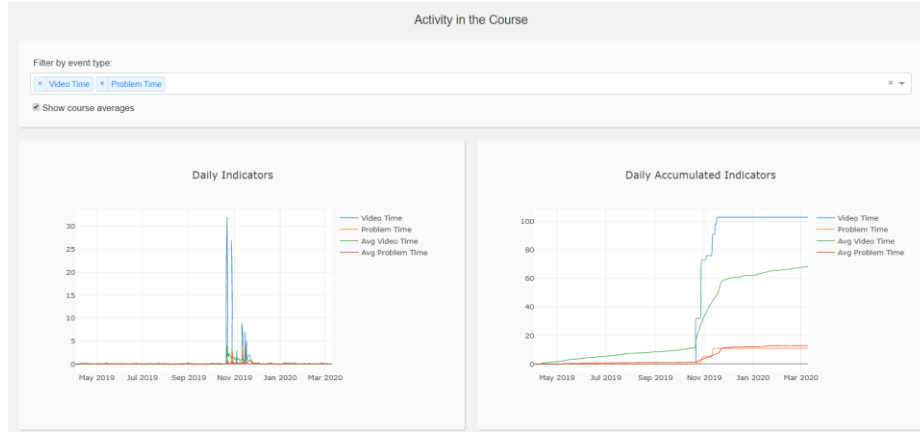


Fig. 3. Graphs with the daily indicators of the learner.

- Data about the performance of the learner (see Fig 4.). The left graph shows the progress of the learner on the content course units and the right graph shows their progress on the graded assignments. In addition, the learner can compare their metrics with the average values of the same ones from all the learners.

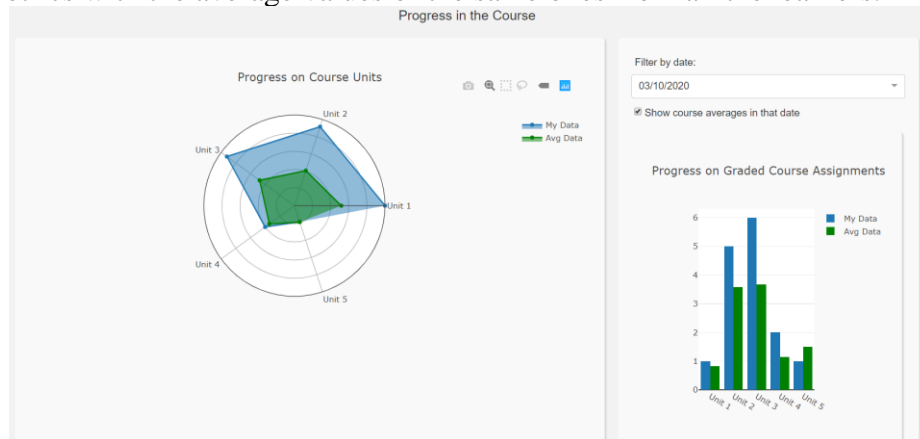


Fig. 4. Graphs with the progress of the learner.

Learner Interest in their Engagement Subservice

The “Intervention Visualisation Service” integrates a functionality that registers all learner interactions with the graphs and their parameters available in their Learner Dashboard in the MongoDB Database. Therefore, this information shall be used in the “Learning Intervention Monitoring Service”.

2.5 Course Data Monitoring Service

This service depicts course data in several graphs and tables as follows:

- General data of the course (see Fig. 5) in these metrics: total number of learners (the learners who are enrolled in the verified itinerary of the course) and certified learners (the verified learners who have passed the course because their grade is equal or greater than 0.5), average grade of learners and average grade of certified learners.

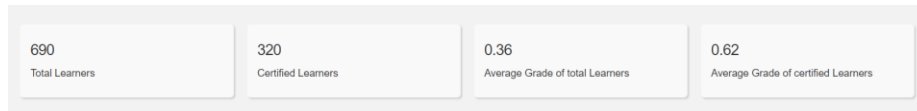


Fig. 5. Table with general course data.

- Demographic data of the learners in the course (see Fig.6) are shown. There are four graphs with the distribution of learners by country, gender, age and academic level.

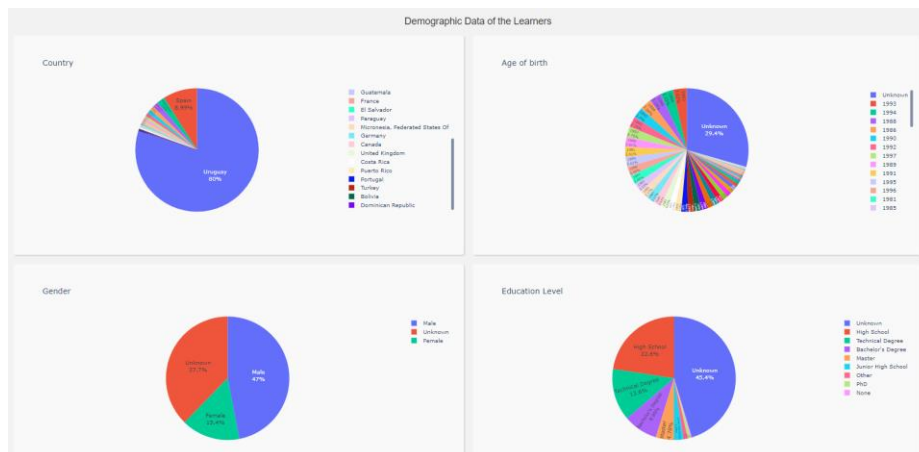


Fig. 6. Graphs with learners' demographic data.

- Data about the learners' interactions in form of their calculated indicators daily (see Fig. 7). It is possible to apply several aggregation operations to these values in the graphs, such as, the average value or total value of the metrics. The left graph shows the selected metrics and the right graph shows these values accumulated in time.

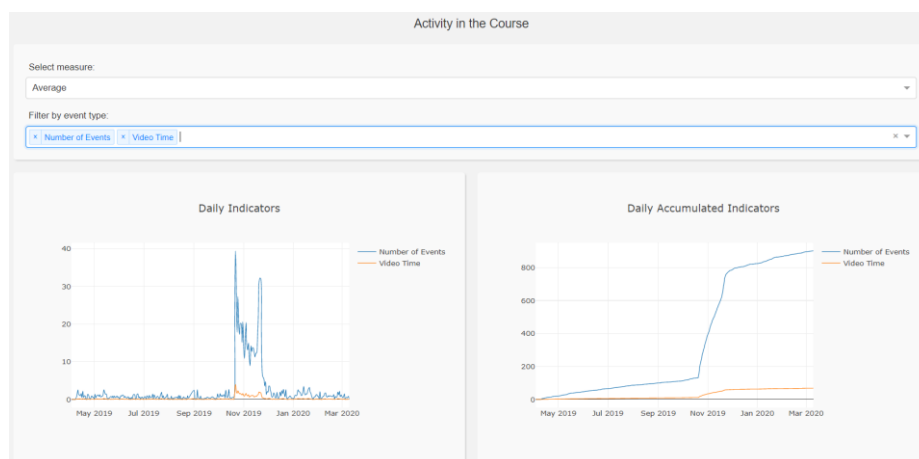


Fig. 7. Graphs with the daily indicators of the learners.

- Data related with the most recent activity of the learners in a table (see Fig. 8) and their grades per unit and in the total course are shown. In addition, it is possible to access to the Learner Dashboard of each learner in the table.

More Data of the Learners and access to their Learner Dashboards

10 results

#	User Id #	Last Login #	Unit 1 #	Unit 2 #	Unit 3 #	Unit 4 #	Unit 5 #	Final Grade #	Learner Dashboard
	filter data...								
	23875584	2020-03-06T19:52:05	1	0.92	0.92	0.4	0.19	0.67	Access
	19517428	2020-02-17T21:28:29	0.5	0.54	0.61	0.75	0.3	0.51	Access
	19522179	2020-03-09T13:08:39	0.75	0.51	0.95	0.75	0.73	0.74	Access
	19505192	2020-03-10T20:21:43	0.75	0.78	0.9	0.5	0.42	0.69	Access
	24051002	2020-01-23T02:01:07	0	0	0	0	0	0	Access
	19461312	2020-03-10T12:49:17	0.75	0.82	0.87	0.75	0.25	0.67	Access
	19467994	2020-02-17T16:07:55	0.5	0.36	0.71	0.8	0.35	0.51	Access
	19480400	2020-01-16T16:15:12	0	0	0	0	0	0	Access
	24072497	2020-01-27T21:33:58	1	0.36	0.68	0.75	0.19	0.48	Access
	19490982	2020-02-04T03:11:48	0.25	0.28	0.82	0.1	0	0.35	Access

Fig. 8. Table with most recent learners' activity and Access to their Learner Dashboards.

2.6 Learning Intervention Monitoring Service

This service allows the monitoring of the learners' interest in the intervention strategy on their learning. It shows the learners' interactions in the different graphs of their Learner Dashboards. It is possible to monitor this issue individually (selecting only one learner, see Fig. 10) or collectively (for all the learners, see Fig. 11). See next section for more details.

3 Case Study at WebApp MOOC

University Autónoma of Madrid offers from 2019 the Massive Open Online Course (MOOC) entitled "Introduction to Development of Web Applications", which is known as "WebApp" (<https://www.edx.org/course/introduccion-al-desarrollo-de-aplicaciones-web-2>). The first run of this course was from February 5th to March 26th 2019. Its second run is in self-paced mode, i.e. everyone can enrol it free from April 9th, 2019. Nowadays, the course is running, and the data presented in this article are obtained in March 2020.

"WebApp" is a five-week course, which contents are structured in five units (one unit per week work). With this course, the learners can learn to develop any Web Application (Web App) based on HTML, CSS, Python, JSON, JavaScript and Ajax. In each course unit, there are multimedia resources, discussion forums and course evaluation activities in form of graded assignments.

We explain the monitoring of the intervention strategy applied to the learning of the learners of the mentioned MOOC with the following case use diagram.

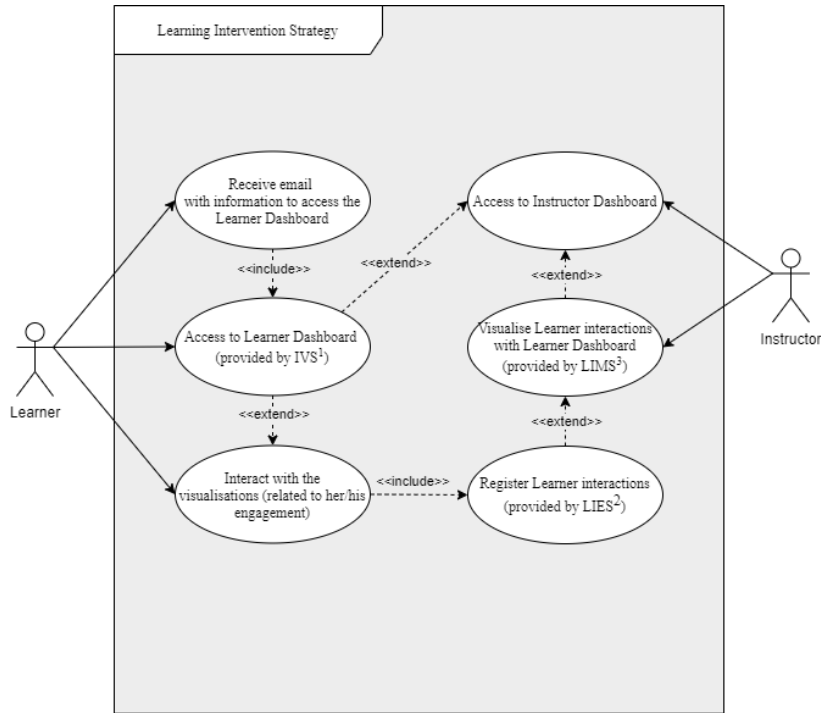


Fig. 9. Case Use Diagram.

1 IVS: Intervention Visualisation Service

2 LIES: Learner Interest in their Engagement Subservice

3 LIMS: Learning Intervention Monitoring Service

Firstly, the instructor can monitor the interest of any learner in their engagement in the course, that is, their interactions with the Learner Dashboard (see Fig. 10). As we can see, the selected learner visited several times their Dashboard in the last week. This learner has a constant interest in their progress in the content course and in the graded assignments.

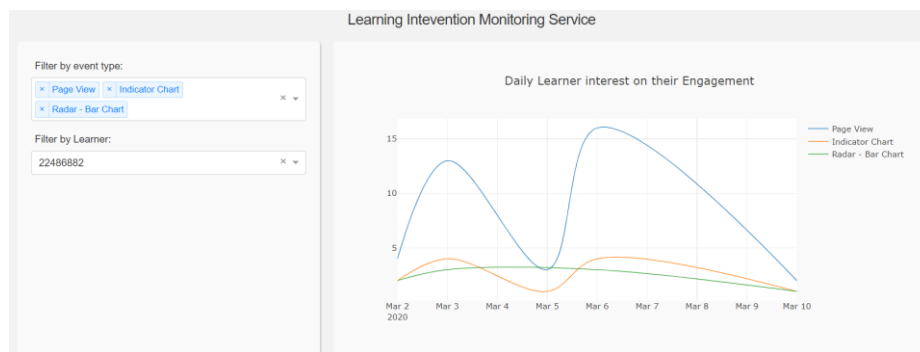


Fig. 10. Graph with the daily interest of a learner on their engagement in the course (individually view).

Secondly, the instructor can monitor the interest of all learner in their engagement in the course, that is, all the learner interactions with the Learner Dashboards are shown (see Fig. 11). As we can see, the maximum numbers of interactions with the dashboard are in two dates separated for a period of one week. This fact is reasonable since the MOOC learners receive weekly the information with the access to their Learner Dashboards. It confirms

learners are interested in this intervention strategy because they use to access to their Dashboards at least weekly.

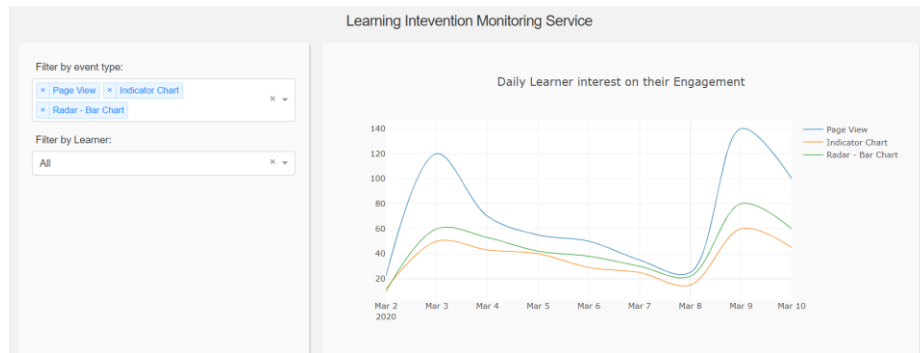


Fig. 11. Graph with the daily interest of all learners on their engagement in the course (collectively view).

4 Conclusions and future work

MOOCs learners are usually affected by feelings of insufficient feedback, lack of interaction with the instructors and isolation feeling while learning, which could lead to dropping out of the course. To mitigate these risk factors, we propose edX-LIMS as a multi-tool Learning Analytics System that facilitates: i) an intervention strategy on learners' learning by visualising their performance and activity in the course in a Learner Dashboard; ii) a graphical monitoring in an Instructor Dashboard that illustrates the interest in this service by the learners and, therefore, facilitates evaluating the success or failure of the aforementioned strategy.

The main contribution of edX-LIMS in comparison with the previous developed system (edX-LIS) is that it gathers about how much is the learner engaged in its course, not only by their grades, but also for the monitoring of its interest on it. With the monitoring of its activity in the Learner Dashboard, the course instructor can easily identify which users are being responsive to the intervention strategy on their learning.

Since the system is independent of edX Platform, there is not a direct synchronization with edX data; hence, it is required to obtain the course data (log tracks) weekly in order to import them to the platform. This is a transmission chain whose disruption could cause the monitoring process to be interrupted, as there would be no new data available to provide feedback to users. This situation is happening now: due to the circumstances caused by COVID-19, we have not been provided with new course data (log tracks) since March, so the system information is currently outdated until new information come to us.

The future development of edX-LIMS aims to increase the quality of its information provided, because the current data information showed on the Learner Dashboard is the result of the answers obtained from a satisfaction questionnaire on the edX-LIS users. Nowadays, we are working in providing a new satisfaction questionnaire for the edX-LIMS users to know how we can improve their experience with the Learner Dashboard and, therefore, we could improve the Learning Intervention Monitoring Service.

ACKNOWLEDGMENTS

This work has been co-funded by the Madrid Regional Government, through the project e-Madrid-CM (S2018/TCS-4307). The e-Madrid-CM project is also co-financed by the Structural Funds (FSE and FEDER). Special thanks are due to the Universidad Autónoma de Madrid for facilitating access to the data of the MOOC used in this article.

References

1. Alaman, X., Carro, R., Cobos, R., Gómez, J., Jurado, F., Molins, P. Montoro, G., Moreno, J., Ortigosa, A., Rodriguez, P., Torrado, JC GHIA: Student Modeling, Learning Analytics, Attention to Diversity and e-Learning. IE Communications. Iberoamerican Journal of Educational Informatics, Number 30, July-December 2019, 1-12.
2. Cobos, R, Gil, S., Lareo, A, & Vargas, F. Open-DLAs: An open dashboard for learning analytics, in: L@S 2016 - Proceedings of the 3rd 2016 ACM Conference on Learning at Scale, ACM, 2016: pp. 265–268. <https://doi.org/10.1145/2876034.2893430>
3. Cobos, R, Jurado, F., Blázquez-Herranz, A. (2019). A Content Analysis System that supports Sentiment Analysis for Subjectivity and Polarity detection in Online Courses. IEEE Iberoamerican Journal of Learning Technologies, 14(4), 177–187. <https://doi.org/10.1109/rita.2019.2952298>
4. Cobos, R., Olmos, L. (2018). A Learning Analytics Tool for Predictive Modeling of Dropout and Certificate Acquisition on MOOCs for Professional Learning, in: 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), IEEE, 2018: pp. 1533–1537. <https://doi.org/10.1109/IEEM.2018.8607541>
5. Davis, D., Chen, G., Jivet, I., Hauff, C., Houben, G.-J. Encouraging Metacognition & Self-Regulation in MOOCs through Increased Learner Feedback, (2016) 17-22 BT-LAL 2016-Learning Analytics for Learn.
6. Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., Rienties, B., Ullmann, T., Vuorikari, R. Research Evidence on the Use of Learning Analytics - Implications for Education Policy, in: R. Vuorikari, J. Castaño Muñoz (Eds.), European Commission's Joint Research Centre Science for Policy Report (JRC), JRC, 2016.
7. Lang, C., Siemens, G., Wise, A., Gasevic D. Handbook of Learning Analytics, SOLAR, Society for Learning Analytics and Research, 2017.
8. Ma, L., Lee, C.S. Investigating the adoption of MOOCs: A technology–user–environment perspective, Journal of Computer Assisted Learning. 35 (2019) 89–98.
9. Martínez-Monés, A., Dimitriadis, Y., Acquila- Natale, E., Álvarez, A., Caerio-Rodríguez, M., Cobos, R., Conde-González, M. A., García-Peñalvo, F. J., Hernández-Leo, D., Menchaca, I., Muñoz-Merino, P. J., Ros, S., y Sancho-Vinuesa, T. (2020). Achievements and challenges in learning analytics in Spain: The view of SNOLA. RIED. Revista Iberoamericana de Educación a Distancia, 23(2), (preprint). doi: <http://dx.doi.org/10.5944/ried.23.2.26541>
10. Romero C, Ventura S. Educational data mining and learning analytics: An updated survey. WIREs Data Mining Knowl Discov. 2020; e1355. <https://doi.org/10.1002/widm.1355>
11. Topali, P., Ortega-Arranz, A., Er, E., Martínez-Monés, A., Villagrà-Sobrino, G.-J., Dimitriadis, Y. Exploring the Problems Experienced by Learners in a MOOC Implementing Active Learning Pedagogies BT - Digital Education: At the MOOC Crossroads Where the Interests of Academia and Business Converge, in: M. Calise, C. Delgado Kloos, J. Reich, J.A. Ruiperez-Valiente, M. Wirsing (Eds.), Springer International Publishing, Cham, 2019: pp. 81–90